(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 September 2001 (07.09.2001)

PCT

(10) International Publication Number
WO 01/64835 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US01/04927

(22) International Filing Date: 26 February 2001 (26.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/515,126	28 February 2000 (28.02.2000)	US
09/577,409	18 May 2000 (18.05.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	09/515,126 (CIP)
Filed on	28 February 2000 (28.02.2000)
US	09/577,409 (CIP)
Filed on	18 May 2000 (18.05.2000)

(71) Applicant *(for all designated States except US)*: **HYSEQ, INC.** [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **LIU, Chenghua** [CN/US]; 1125 Ranchero Way #14, San Jose, CA 95117 (US). **DRMANAC, Radoje, T.** [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agents: **ELRIFI, Ivor, R.** et al.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

WO 01/64835 A2

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

Technology aimed at the discovery of protein factors (including *e.g.*, cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (*i.e.*, partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NOS: 1-13901. The polypeptides sequences are designated SEQ ID NOS: 13902-27802. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO: 1-13901 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NOS: 1-13901. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO: 1-13901 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1-13901. The sequence information can be a segment of any one of SEQ ID NO: 1-13901 that uniquely identifies or represents the sequence information of SEQ ID NOS: 1-13901.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-13901 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-13901 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO: 1-13901; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NOS: 1-13901; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NOS: 1-13901. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO: 1-13901; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (*e.g.*, SEQ ID NOS: 13902-27802); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (*e.g.* orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO: 1-13901 ; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (*e.g.*, with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (*e.g.* host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, *e.g., in situ* hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (*i.e.*, increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (*e.g.*, bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs: 1-13901 .

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NOS: 1-13901. The sequence information can be a segment of any one of SEQ ID NO: 1-13901 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NOS: 1-13901. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4^{20} possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match ($1 \div 4^{25}$) times the increased probability for mismatch at each nucleotide position (3 x 25). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements *e.g.* repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (*e.g.*, with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant"(or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, *e g.*, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (*e.g.*, nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (*e.g.*, microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (*e.g.*, yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g.*, *E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (*e.g.*, soluble proteins) or partially (*e.g.*, receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (*e.g.* Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134 -143) and factors released from damaged cells (*e.g.* Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (*i.e.*, hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (*i.e.*, washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (*i.e.*, the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, *e.g.*, mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, *e.g.*, mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (*e.g.*, via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, *e.g.*, using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, *e.g.* by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO: 1-13901 ; a polynucleotide encoding any one of the peptide sequences of SEQ ID NOS: 13902-27802; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NOS: 13902-27802. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO: 1-13901; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NOS: 13902-27802. Domains of interest may depend on the nature of the encoded polypeptide; *e.g.*, domains in receptor-like polypeptides include ligand-binding,

extracellular, transmembrane, or cytoplasmic domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, *e.g.*, cDNA and genomic DNA, and RNA, *e.g.*, mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO: 1-13901 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO: 1-13901 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO: 1-13901 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, *e.g.*, at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO: 1-13901, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, *e.g.* 15, 17, or 20 nucleotides or more that

are selective for (*i.e.* specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO: 1-13901, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO: 1-13901 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO: 1-13901 can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative choices (*e.g.*, hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO: 1-13901 , or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, *e.g.*, plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-13901 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-13901 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and

promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

4.3 ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1-13901 , or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

NOS: 13902-27802 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO: 1-13901 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (*e.g.*, SEQ ID NO: 1-13901), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a n omeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -uni ts, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (*i.e.*, SEQ ID NO: 1-13901). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (*e.g.*, promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, Proc. *Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (*e.g.*, by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5′ flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NOS: 13902-27802 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO: 1-13901 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO: 1-13901 or

(b) polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NOS: 13902-27802 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NOS: 13902-27802 or the corresponding full length or mature protein; and "substantial equivalents" thereof (*e.g.*, with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NOS: 13902-27802.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, *e.g.*, Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: *A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, *e.g.*, ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NOS: 13902-27802.

The protein of the invention may also be expressed as a product of transgenic animals, *e.g.*, as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, *e.g.*, U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl™ or Cibacrom blue 3GA Sepharose™; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP- HPLC) steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (*i.e.*, glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e,g.,* cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (*e.g.*, liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (*e.g.*, by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, *e.g.*, homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, *e.g.*, via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells *in vivo* or *ex vivo* is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, *e.g.* in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (*i.e.*, traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (*i.e.*, in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), *e.g.*, in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (*e.g.*, HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (*e.g.*, anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (*e.g.*, a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (*e.g.* proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (*e.g.*, stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, *e.g.* reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (*e.g.* exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, *e.g.* from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1- 7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14 . Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (*i.e.*, increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science 282*:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol*, 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol*, 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem*, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, *e.g.*, ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide *e.g.* a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (*i.e.*, increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

(i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;

(ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

(iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;

(iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;

(v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;

(vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

(viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

(ii) increased sprouting of neurons in culture or in vivo;

(iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or

(iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, *etc.*, depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, *e.g.*, weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape);

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

4.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, *e.g.*, differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, *e.g.*, by an antibody specific to the variant sequence.

4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01μg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1μg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF-α and TGF-β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (*e.g.*, heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (*e.g.*, at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, *e.g.*, treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, *e.g.* polyvinyl pyrrolidone; and other

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μg to about 100 mg (preferably about 0.1 μg to about 10 mg, more preferably about 0.1 μg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, *e.g.*, amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50}. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, *e.g.*, Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μg/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μg/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab}, $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG_1, IgG_2, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

5.13.1 Polyclonal Antibodies

For the production of polyclonal antibodies, various suitable host animals (*e.g.*, rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (*e.g.*, aluminum hydroxide), surface active substances (*e.g.*, lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

5.13.2 Monoclonal Antibodies

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (*e.g.*, by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5.13.2 Humanized Antibodies

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, *e.g.*, mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al,(Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse™ as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

5.13.4 F_{ab} Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see *e.g.*, Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

5.13.5 Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (*e.g.* tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (*e.g.* alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (*e.g.* $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

a leukocyte such as a T-cell receptor molecule (*e.g.* CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc R I (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

5.13.6 Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (*e.g.*, avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (*e.g.* text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO: 1-13901 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO: 1-13901 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (*e.g.*, where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, *e.g.*, Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

4.18 SCREENING ASSAYS

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO: 1-13901 , or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

(a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and

(b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

4.19 USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NOS: 1-13901. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO: 1-13901 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, *i.e.*, small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.*, (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-$MeIm_7$), is then added to a final concentration of 10 mM 1-$MeIm_7$. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-$MeIm_7$, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, *e.g.*, Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *Cvi*JI, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease *Cvi*JI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (*Cvi*JI**), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *Cvi*JI** digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that *Cvi*JI** restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm^2, depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm^2 and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers *e.g.* a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

5.0 EXAMPLES

5.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (*e.g.*, 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

5.2 EXAMPLE 2

Novel Contigs

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-13901 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (*i.e.*, Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO: 1-13901) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: : 1-13901. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from http://fasta.bioch.virginia.edu) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-13901 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database release 200101 (Derwent), using BLAST algorithm. The nearest neighbor result showed the closest

homologue for SEQ ID NOS: 1-13901. The nearest neighbor results for SEQ ID NO: 1-13901 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NOS: 1-13901. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/515,126

TABLE 1

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
adult brain	GIBCO	AB3001	83 544 597-598 600-607 616 841 1004 1148 1346 1493 1974 2138 2141 2143 2161 2266 2345 2363 2511 2569 2876 2880 3001 3099-3101 3105-3106 3110-3111 3115-3117 3199 3272 3282 3284 3356 3425 3537 3634 3689 3709 3797 3810 3839 3899 4006 4021-4022 4025 4043 4194 4201 4253 4277 4297 4388 4399 4410 4667 4671 4722 4747-4748 4750 4755 4767 4845 4865 4940 5037 5075 5093 5118 5163 5171-5172 5268 5481 5523 5553 5656 5724 5894 5902 5938 6052 6170-6173 6176 6214 6307 6336 6369 6374 6793 6894-6897 6979 7058 7169 7455 7492-7493 7495-7499 7501 7504 7577 7586 7761 7792 7864 7870 8035 8065 8085 8110 8120 8140 8224 8226 8298 8372 8427 8452 8456 8535 8648 8672 8674-8679 8681-8684 8816 8838-8839 8870 8898 9012 9041 9079 9128 9257 9264 9304 9317 9460 9503 9517 9567 9623 9734 9781 9792-9798 9929 9964 9999 10296 10330 10469-10470 10578 10679 10778 10786 10895 10984-10986 11032 11052 11069 11130 11145 11239 11289 11402 11818 11862 11870-11876 11878-11881 12017 12037 12127 12160 12294 12363 12375 12405 12424 12438 12467 12539 12570 12590 12615-12616 12618 12685 12688 12712 12739 12748 12830 12913 12916 12948-12950 13002 13064 13073 13083 13141 13150 13153 13164-13166 13257 13391 13456 13479 13489 13492 13494 13499 13501 13503 13560 13595-13596 13627 13645 13679 13782 13795 13861 13866 13869 13882
adult brain	GIBCO	ABD003	67 83 142 443 587 598 608-609 611 613-624 633 731 734 737-742 760 799-800 809 1148 1152 1167-1184 1193 1346 1433 1516 1552 1575 1671 1756 1774 1833 1974 2138 2145 2176-2178 2237 2266 2299-2301 2303-2306 2343 2363 2412 2444 2449 2511 2516 2555 2569 2576 2614 2716 2809 2876 2911 2926 3001 3093 3114 3119 3121-3124 3126 3128-3130 3234 3254-3256 3258-3263 3265-3267 3270-3274 3276-3277 3280-3281 3284 3286 3348 3356 3378 3435 3459 3484 3537 3548 3595 3605 3625 3627 3634 3686-3697 3700 3702 3709 3711 3720 3722 3737 3757 3797 3804 3810 3839 3856 4006 4019 4025 4040 4055 4057-4058 4060 4078 4194 4201 4246 4253 4277 4282 4390 4405 4412 4431 4620 4622 4641 4689 4751-4764 4791 4808 4837 4845 4847-4849 4852-4858 4860-4862 4864-4869 4940 4957 4962 4972 4998 5021 5031 5037-5038 5040 5076 5093 5108 5118 5167 5169 5171-5172 5251-5261 5263-5265 5270 5364 5401 5481 5492 5521 5523 5535 5656 5674 5693 5766 5788 5817 5906-5909 5938 6005 6027 6057 6064 6147 6178 6180-6182 6189 6214 6229-6233 6254 6272 6369 6371 6421-6426 6555 6595 6598 6601 6799 6803 6825 6836 6886 6894 6913 6972 6995 7058 7104 7130 7133 7148 7164 7169 7339 7347 7386 7426 7455 7494 7502 7507 7509 7511-7512 7516 7520 7584-7587 7590-7596 7598-7601 7603-7604 7608 7632 7677 7743 7748 7761 7768 7792 7797 7807 7815 7839 7849-7861 7864 7870 7930 7937 8035 8065 8067 8080 8087 8095 8110 8120 8139-8140 8209 8224 8226 8235 8246 8262 8285 8298 8320 8323 8336 8354 8361 8365 8370 8375 8387 8452 8456 8535 8556 8576-8577 8603 8630 8648 8674 8685-8686 8688-8690 8693 8695 8702 8712 8742 8760-8761 8763-8764 8766-8769 8813 8815-8816 8830 8834 8838-8839 8848 8863 8870 8898 8921 8943-8944 8951 8989 9010 9041 9050-9056 9058-9064 9076 9079 9092 9097 9128 9144-9145 9257 9264 9271 9278-9279 9304 9315 9317 9455 9466 9472 9475 9480 9503 9511 9517 9525 9539 9689 9734 9773 9781 9791 9799-9802 9847 9852 9873 9928-9929 9964 9999-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			10001 10003-10004 10053 10175 10267 10276 10330 10349 10434 10449 10466 10471-10474 10492 10496 10509 10566 10578 10597-10599 10645 10679 10788 10891 10948 10988-10990 11032 11039 11041-11043 11052 11065 11069 11105 11108 11130 11145 11167 11196 11203-11207 11209 11239 11399 11401-11402 11406 11459 11470 11604 11606 11642 11761 11818 11862 11877 11882-11884 11886 11889-11893 11944-11946 11981 11988 12016 12019 12022 12037 12083 12127 12143 12164-12165 12168-12171 12178 12195 12236 12265 12305 12327 12363 12375 12405 12423-12424 12430 12438 12546 12570 12590 12594 12612 12615-12618 12630 12670 12674 12685-12688 12693 12704 12706-12707 12748 12772 12830 12885 12904 12913 12916 12923 12933 12951 12956 12993 13001 13020-13021 13038 13047 13064 13072-13073 13084-13085 13092 13117 13142 13167-13171 13191 13254 13257 13260 13295 13390-13391 13394 13456 13479 13483 13489 13497 13501 13503 13505-13507 13512 13516 13546 13551 13555 13575 13590 13592 13597 13613-13614 13645 13649 13659 13711 13782 13795 13838 13861 13869 13875 13882 13884-13885 13888 13892 13896
adult brain	Clontech	ABR001	142 858 1542 2174 2407 2483 2652 3272 3287 3460 3492 3535 3595 3737 3839 4005 4060 4282 4434 4791 4972 5040 5293 5523 5530 5535 5788 5906 6082 6601 6799 6980 7373 7577 7587 7759 7788 7851 8081-8082 8110 9167 9455 9466 9781 9928 10422 10774 10791 11069 11401 11406 11459 11604 11607 11791 11818 11865 11961 11979 12022 12122 12160 12327 12442 12594 12615 12640 12670 12705 12935 12957 12985 13047 13197 13257 13456 13511-13512 13546 13554 13646 13793 13885 13889 13893
adult brain	Clontech	ABR006	6 67 1004 1908 3272 3286 3548 4011 4282 4998 5923 5928 6374 6730 6815 6867 6890 7067 8365 9264 9729 9780 10776 11587 11618 12596 12601 12605 12704 12749 12754 12951 13047 13051 13090 13479 13488 13498-13499 13503 13512 13575 13882
adult brain	Clontech	ABR008	6 11 21 41 51 88 142 364 376 579 598 651 736 800 1050 1148 1184 1251-1265 1291 1346 1404 1479 1529 1543 1671-1674 1697 1699-1710 1820 1830 1832-1838 1840 1848-1849 1908 1914 1919 1927 1957 1964 1974 1976 1978-1979 2005-2006 2050 2081 2090 2110-2111 2129 2150 2174 2200 2310 2327 2342 2408-2410 2420 2444 2449 2461-2467 2484 2490 2499-2506 2511 2553 2574 2576 2611 2652 2809 2827 2866 2894 3032 3207 3535 3591 3610 3634 3715 3722 3737 3766-3770 3819 4006-4007 4011 4025 4032 4060 4078 4095 4109 4128 4143-4155 4182 4194 4247-4257 4277 4282 4294 4296 4310 4330 4348 4355 4360 4381 4395 4399 4411 4431 4543 4641 4662 4694 4698 4767 4781 4791 4808 4833 4837 4985 5001 5022 5040 5075 5094 5108 5163 5303-5306 5308-5314 5320 5380 5523 5553 5615 5625-5626 5634 5638-5644 5701 5706 5711 5727-5742 5766 5772 5775 5783 5801 5814 5817 5820-5821 5829 5837 5851 5855 5858 5864 5867 5874 5885 5890 5897 5901 5906 5923 6057 6125 6214 6223 6288 6302 6456-6464 6545 6598 6601-6602 6624 6676-6685 6699 6726 6728 6746-6749 6765 6799 6805 6854 6860 6893-6894 7004-7007 7049 7076 7078 7081-7083 7105 7117 7119 7133 7153 7166 7431 7579 7708 7768 7849 7900-7905 8018 8083-8084 8095 8110 8196-8208 8262 8288 8312 8320 8331 8336 8356 8375 8452 8482 8633 8681 8710 8739 8777 8815 8817 8830 8839 8963 8965 8983 9010 9097 9100 9102-9108 9111 9128 9142 9257 9264 9313 9364 9378-9384 9401 9454-9455 9458 9460-9464 9503 9509 9511 9515-9516 9522 9528-9529 9533 9539 9542 9544 9573 9577 9646 9773 9780 9924 10000 10025-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			10029 10097 10148 10209 10218-10223 10225-10229 10267 10275-10276 10284 10292 10296 10303 10313 10326 10329-10331 10333-10334 10343 10346 10393 10430 10494 10496 10513 10542 10557 10613-10616 10679 10688 10691 10719-10721 10729 10743 10758 10760 10776 10782 10786 10795 10895 11100 11114 11132 11195 11240 11242-11254 11401 11406 11437 11454 11459 11462 11466-11473 11475-11476 11519 11532-11535 11550 11553 11555-11556 11559 11588-11589 11606 11615 11618 11621 11627-11628 11633 11761 11791 11807 11818 11932 11955 12006 12039 12041 12092 12212-12224 12231 12236 12305 12363 12368 12396-12399 12405 12424 12439 12442 12465-12466 12468 12477 12479-12482 12490 12521 12546 12552-12553 12576-12577 12579 12582 12585 12590 12601 12609 12617-12618 12636 12658 12707 12725 12735-12736 12749 12754 12776-12779 12859 12868 12894 12905 12909 12940 12955 12959 12977 12980 12990 13002 13004-13005 13020 13035 13038 13042-13044 13047 13051-13052 13056 13062 13073 13082-13083 13196 13249 13280 13311-13312 13336 13387 13417 13421-13426 13436 13445 13456 13458 13479-13482 13488 13490 13494-13495 13497-13500 13503 13507 13512 13516 13533 13546 13554-13555 13590 13613 13630 13649 13659 13670 13678 13713 13724 13769 13793-13794 13808 13827-13828 13838 13861 13867-13868 13875 13882 13884-13885 13888-13889 13893 13896 13898
adult brain	Clontech	ABR011	1006 1257 3797 4006 4025 5535 6057 7169 7870 8262 8937 8966 9257 10778 12736 13394 13679 13793 13861
adult brain	BioChain	ABR012	88 598 1007 1134 2597 3557 3590 3627 3797 4006 4192 4246 4282 4391 4940 5523 5535 6288 6338 7138 8110 8898 9076 9401 9455 9476 10772 11061 11114 12989 13394 13511 13866
adult brain	Invitrogen	ABR013	598 2614 3191 4355 4391 5523 5788 8085 8486 11513 12521 12989 13861
adult brain	Invitrogen	ABT004	40 51 598 1050-1057 1148 1777-1778 1947 1976 2270-2272 2327 2490 2617 3050 3600-3602 3722 3987 4390-4391 4434 4543 4689 5031 5157-5159 5167 5169 5466 5505 5682-5683 5701 5766 5778 5794 5902 6147 6367-6371 6459 6545 6709 6728 6783 6801 6971 7104 7175 7815 7839 7864 8139 8342 8345 8355 8363 8372 8452 8633 8963 8975-8976 9012 9133 9423-9424 9511 9515 9517 9528 9556 9827 9949 10260 10267 10275 10570-10571 10733 10767 11132 11159 11406 11459 11932 12009 12092 12109-12111 12127 12283 12428 12511 12579 12605 12725 12747 12830 12885-12886 12910 12913 12954 12987-12989 13051 13054 13062 13073 13090 13249-13253 13438 13445 13456 13489 13500 13512 13516 13533 13546 13590 13622 13649 13683-13684 13713 13803 13838 13861 13866 13896
cultured preadipocytes	Strategene	ADP001	1134 1346 2343 2614 3272 3426 3610 3720 3839 3885 4011 4277 4282 4297 4346 4388 4391 4405 4434 4641 4833 4940 4985 5018 5030 5040 5163 5167 5523 5581 5778 5788 5794 5895 5951 6082 6147 6272 6607 7067 7141 8093 8235 8285 8312 8363 8629 8648 8830 8839 9290 9401 9466 9503 9781 10346 10470 10776 10795 10971 11108 11170 11513 11818 12034 12037 12046 12093 12375 12387 12405 12424 12570 12636 12670 12674 12688 12735 12749 12913 12940 13126 13163 13295 13489 13494 13497 13499 13511 13516 13575 13652-13653 13866 13888-13889
adrenal gland	Clontech	ADR002	8 83 142 225 351 443 551 569 731 864 1134 1266-1271 1273-1274 1276-1292 1294-1295 1381 1391 1544-1545 1658 1671 1908 1959 1983 2010 2023 2145 2175 2283 2310 2328-2334 2343 2444 2449 2510 2522 2576 3032 3069 3153 3166 3272

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			3378 3416 3548 3625 3709 3711 3771-3788 3790-3791 3797 3839 3870 3899 3985 4019 4054 4060 4109 4173 4192 4194 4201 4253 4277 4282 4389 4391 4395-4396 4431 4667 4687 4694 4783 4957 4966 4985 4998 5037 5108 5293 5316-5323 5325-5336 5481 5505 5527 5531-5533 5560 5628 5656 5701 5766 5865 5874 5902 5939 5979 6020 6052 6057 6227 6288 6354 6405 6449 6465-6482 6521 6603 6726 6894 6979 7008-7011 7155-7156 7204 7604 7634 7845 7871 7906-7912 7915-7918 7920-7930 8022 8067 8085-8086 8095 8110 8116 8224 8262 8363 8365 8412 8520 8535 8554 8699 8742 8831 8870 8950-8951 9002 9010 9012 9097 9109-9119 9121-9127 9190 9264 9280 9304 9317 9455 9457 9495 9503 9553 9556 9586 9709 9780-9781 9895 9927 10007 10030-10032 10034-10035 10037-10045 10119 10277 10284 10346 10595 10617-10620 10623-10627 10645 10675 10679 10760-10761 10766 10774 10782 11108 11194 11255 11258-11270 11289 11400 11406 11457 11519 11588 11600 11621 11626 11818 11952 12080 12159 12225 12227-12230 12232-12237 12239-12240 12242-12249 12251 12305 12363 12375 12404-12405 12424 12439 12577 12599 12601 12630 12636 12657-12658 12663 12688 12693 12706 12713 12715 12735 12749 12754 12780-12787 12848 12863 12909 12913 12923 12957 12990 13006-13007 13020-13021 13064 13095 13104 13295 13313 13315 13388-13389 13391 13456 13473 13494 13507 13515 13546 13613 13631 13679 13725-13733 13858 13866 13872 13883-13884 13888-13889
adult heart	GIBCO	AHR001	51 83 88 94 221 239 360 366-367 404 410-411 413 415 458-459 461 465-468 471 473-478 486 545-546 559 567 616 625-630 743-744 799 802-806 808 810-835 837-842 959 1004 1066-1085 1134 1178 1184-1193 1346 1512 1516 1546-1547 1556 1575 1671 1727 1774 1829 1959 1976-1977 2090 2096 2108 2110 2128 2138 2145-2147 2161 2179 2195-2198 2257 2276 2278-2281 2302 2307-2309 2363 2398-2399 2409 2411-2412 2444 2449 2497 2516 2529 2563 2569 2575-2576 2597 2605 2614 2617 2762 2809 2816 2879-2880 2911-2924 2926 2961 2978-2980 2985-2986 2993 2995-3002 3032 3042 3051 3058 3069 3081 3091-3094 3109 3114 3132-3133 3135 3137-3138 3141 3191 3196 3199 3215 3263 3272 3282-3286 3317 3349-3350 3353-3359 3361-3362 3364-3367 3370-3393 3396-3399 3403 3406 3425 3451 3465-3466 3479 3500 3503 3537-3538 3544 3548 3550 3555 3557 3590-3591 3595 3604 3606-3612 3614-3621 3623-3627 3634 3689 3697-3698 3701-3709 3711-3713 3720 3722 3737 3757 3797 3839 3885 3898 3988-3989 3996 4005-4006 4008 4010-4011 4019 4021-4022 4025 4040 4043 4054-4055 4058-4060 4078 4183 4192 4194 4201 4246 4253 4269 4277 4282 4341 4351 4391 4403 4405 4434 4517 4543 4553 4590 4615 4622 4631 4633-4634 4641 4654 4664 4666-4667 4672-4675 4688-4689 4693-4694 4738 4740 4755 4783 4828 4870-4871 4905-4906 4909-4915 4917-4919 4921-4924 4926-4927 4940 4957 4972 4985 4991 4998 5001 5030 5037-5038 5040 5076 5118 5163 5171-5175 5177-5178 5180-5182 5266-5268 5380 5422 5481 5490 5492 5506-5507 5523 5534-5535 5581 5656 5682 5711 5740 5766 5788 5895 5908 5923-5924 5938 6047 6057 6083 6085 6104 6110-6111 6117 6147 6184-6185 6189 6195 6215 6255-6257 6259-6266 6268 6272 6288 6307 6336 6375-6379 6382 6427-6429 6458 6555 6588 6595 6607 6629 6665 6689 6765 6767 6799 6815 6871-6873 6879 6881-6882 6898 6926-6928 6975-6978 6987 7058 7090 7148 7313-7314 7341 7343-7344 7346-7347 7350 7354 7359 7362 7373 7380 7394 7402 7407-7408 7410 7413 7415-7416 7418-7419 7426 7431 7468 7473 7480 7494 7505 7513-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			7516 7520 7544 7561 7584 7587 7599 7601-7604 7635 7638-7643 7645-7649 7651-7655 7657 7659-7662 7733 7735 7743 7748 7783-7796 7815 7852 7857 7862-7863 7865-7870 7930 7933 7983 8062-8065 8067 8087-8088 8093 8095 8110 8116 8120 8139-8140 8224 8226 8235 8262 8298 8336 8344-8345 8354 8356 8363 8368 8372 8378 8387 8410 8427 8452 8456 8531-8532 8534-8535 8563-8569 8572 8576-8577 8592-8593 8597 8603 8606 8610 8613-8614 8616-8617 8646 8648 8670 8681 8691 8698-8699 8702 8712 8742 8756 8760 8763 8795 8807-8809 8811-8814 8816-8819 8821-8833 8835-8839 8858 8863 8870 8898 8921 8927 8936 8939 8943 8946 8950-8951 8966 8988-8992 8994-8995 9010 9049 9059 9065-9067 9070-9072 9076 9097-9098 9167 9190 9257-9260 9262 9269 9281 9287 9301 9304 9315 9317 9401 9454-9455 9466 9476 9480 9484 9556 9577 9612 9689 9698 9720-9721 9734 9741-9743 9747-9750 9758 9781 9791 9804 9871-9882 9884-9885 9928 9939 9942 9954-9960 9999-10000 10005 10175 10179 10275 10284 10292 10296 10329-10331 10346 10400 10422 10430-10431 10437 10442 10444-10447 10475 10511-10512 10514-10521 10557 10576-10577 10616 10645 10679 10691 10729 10742 10744 10772 10774 10777-10778 10782 10788 10839 10891 10894-10895 10902 10917-10920 10937 10942 10946 10948 10969-10970 10992 11032 11044-11045 11061 11066-11074 11108 11114 11132 11145 11153 11165-11170 11173 11205 11208 11210-11215 11283 11289 11386 11388 11401-11402 11406 11462 11559 11565 11576 11596 11606 11615 11620 11744 11788-11789 11793 11818 11823-11825 11831-11832 11869 11894 11947 11961 11982-11989 12000 12006 12009 12011 12019-12020 12028 12037 12044 12078 12081 12093 12119-12122 12143 12160 12166 12172-12175 12177-12179 12197 12305 12335 12363 12375-12376 12383 12387 12400 12402 12405 12424 12428 12438 12479 12521 12523 12546 12560 12564 12570 12590 12599 12601 12605 12609 12611 12616-12618 12653 12662-12663 12670 12674 12688 12717-12720 12734 12752 12754 12772 12905 12907 12914 12916-12917 12923 12925 12940 12961 12963-12965 12989 12991 12993 13020-13021 13033 13035 13072-13073 13082-13083 13104 13117 13126 13132 13136 13141-13142 13148 13169 13203-13210 13212 13260-13261 13293 13295 13326 13377 13394 13413 13442 13456 13477 13480 13488 13490 13494-13496 13501-13503 13506-13507 13511 13516 13533 13568 13570-13571 13575 13582 13592 13613 13624-13630 13632 13644 13646 13659-13660 13678-13679 13689 13701 13711-13713 13775 13782 13795 13797 13866-13869 13872 13882 13884-13885 13893
adult kidney	GIBCO	AKD001	49 67 83-84 142 354 405-407 415-429 431-432 445 460 462 479-484 486 488 492-493 524 548-549 598 616 631-638 744 787 809 841 1004 1068 1086 1160 1163-1166 1171 1184 1193 1346 1359 1449 1479 1516 1552 1556 1671 1724 1727 1774 1826 1858 1914 1974 1976 1978-1979 2081 2097-2099 2111-2113 2117-2118 2129-2130 2138 2145 2148-2150 2161 2186 2218 2266 2268 2302 2310 2327 2343 2363 2409 2412 2444 2449 2468 2483 2523 2569 2576 2614 2617 2827 2845 2876 2910 2915 2926-2931 2933-2934 2938-2943 2945-2947 2955-2956 2976 2981 2983-2984 3001-3016 3018 3053 3109 3114 3140 3142-3146 3149-3150 3199 3254 3265 3272 3283-3284 3286 3333 3350 3356 3378 3406 3435 3445 3460 3492 3503 3535 3537 3544 3548 3590-3591 3598 3625 3627 3634 3676-3677 3679-3685 3689 3697 3709 3711 3720 3722 3737 3757 3797 3808 3810 3839 3885 3989 4005-4006 4011 4019 4022 4025 4040 4043 4054-4055 4060 4078 4109 4192 4194 4201

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			4246 4253 4269 4277 4341 4354 4387-4388 4390 4398 4402 4405 4410 4434 4473 4540 4543 4583 4591-4592 4607 4637 4640-4641 4649 4654 4665 4668 4676-4677 4680-4685 4688-4689 4711-4712 4751 4758 4766-4768 4783 4808 4833 4836-4837 4845 4874 4904 4940 4957 4962-4963 4972 4983 4991 4998 5022 5037 5040 5076 5093 5118 5143 5163 5171-5172 5246-5249 5281 5362 5364 5380 5422 5481 5521 5523 5526-5527 5535 5656 5693 5726 5740 5766 5769 5778 5788 5794 5894 5902 5908 5911 5916-5917 5923 5928 5938 5990 6005 6049-6051 6057 6078 6082 6086 6088 6105-6107 6112-6117 6120-6122 6147 6159 6186-6189 6215 6257 6272 6288 6369 6418 6423 6430 6555 6595 6598 6738 6743 6765 6767 6793 6799 6805 6815 6836 6857 6867-6870 6872-6873 6878 6884 6886 6893 6899-6900 6931 6947 6976 6979 6988 7024 7045 7058 7104 7118 7155 7275 7281 7306 7312 7338 7347-7348 7351-7352 7355 7357 7359-7363 7366 7369 7371-7373 7381 7383 7386 7400 7402 7404-7406 7408 7420 7422 7424 7426-7428 7430-7431 7433-7435 7439 7455 7465-7468 7494 7502 7506 7517-7521 7579 7587 7604 7634 7639 7642 7733 7735 7743 7748 7792 7797 7839 7842-7846 7848 7852 7857 7862 7864 7870 7930 7933 8065 8077 8084 8087 8093 8095 8105 8110 8116 8120 8139-8140 8156 8224 8226 8235 8262 8320 8336 8345 8351 8354 8359 8364 8368 8370 8372 8375 8377 8387 8427 8452 8456 8461 8486 8518 8520 8527 8535 8563-8564 8566 8572-8573 8575-8577 8580-8583 8585 8588-8589 8597 8603 8618-8621 8623 8625-8626 8628 8630 8634 8647-8650 8664 8674 8677-8678 8691 8694 8701-8704 8706 8711 8722 8740 8742 8763 8792 8798 8806 8812-8813 8816 8830 8838-8839 8848 8853 8863 8876 8898 8921 8935 8943-8944 8948 8951 8966 8989 9012 9041 9047-9049 9076 9092 9128 9137 9167 9218 9244 9249 9257 9264 9287 9304 9315 9317 9424 9455-9457 9466 9472 9475 9484 9503 9511 9517 9529 9536 9542 9544 9558 9612 9683 9699 9703 9722-9723 9725 9734 9744 9752-9755 9758 9772-9773 9780-9781 9805-9807 9841 9843 9927-9929 9939 9942 9969 9998-10000 10007 10175 10275 10284 10287 10292 10319 10346 10376 10423-10424 10431 10442 10446 10448 10450 10475 10480 10496 10542 10557 10645 10679 10774 10778 10782 10788 10895 10921-10928 10947 10949-10953 10986 10994-10996 11032 11052 11061 11069 11114 11130 11132 11143 11145 11197-11202 11205 11208 11212 11239 11289 11401-11402 11406 11459 11513 11586 11596 11604 11607 11618 11620 11695 11711 11759 11761 11790-11791 11793-11794 11818 11820-11822 11834-11836 11857 11865 11869 11895-11899 11950 11970 12006 12022 12041 12078 12143 12159-12160 12178 12195 12197 12256 12265 12305 12327 12359 12363 12375 12387 12405 12424 12428 12438 12442 12467 12511 12521 12523 12535 12539 12546 12564 12570 12590 12599 12601 12609-12611 12616 12618 12630 12653 12656 12664 12666-12667 12670 12674 12688-12689 12691 12739 12754 12830 12834 12904 12913-12914 12916-12917 12923 12940-12941 12951-12952 12956 12972 12976-12977 12993 12999 13002 13052 13062 13064 13066 13072-13073 13082-13083 13095 13104 13127 13131 13133 13136 13141 13143-13146 13148 13154-13155 13169 13172 13184 13195-13196 13249 13254 13260 13263 13280 13295 13307 13319 13394 13442 13456 13477 13479-13480 13488-13490 13492 13494-13495 13497-13503 13506-13507 13512 13515-13516 13546 13551 13554-13556 13572-13574 13590 13592 13613 13627 13631 13644-13645 13656 13659-13660 13665 13670 13710 13713 13776 13782 13795 13859-13860 13864 13866-13868 13872 13882-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			13885 13888 13891-13892
adult kidney	Invitrogen	AKT002	67 142 550-552 598 639-641 1004 1015 1493 1516 1947 2215 2299 2343 2363 2449 2618 3054-3055 3152-3153 3155-3157 3434 3535 3590 3709 3797 3808 3839 3885 4011 4022 4025 4040 4253 4277 4391 4405 4671 4759 4767 4769 4837 4949 4972 5001 5035 5037 5052 5108 5526 5581 5615 5726 5788 5895 6062 6139 6190-6191 6803 6900 6947 6975 7468-7469 7473 7733 8095 8110 8139 8262 8323 8361 8363 8375 8520 8539 8648 8711 8798 8912 8950 8966 8983 9076 9264 9368 9510 9517 9665 9703 9734 10175 10476 10791 10989 10997 11132 11618 11745 11900 12006 12039 12160 12363 12375 12405 12424 12685 12702 12707 12904 13035 13060 13104 13136 13295 13394 13456 13488 13495 13501 13512 13533 13554 13583 13644 13670 13679 13696 13713 13795 13866 13888 13891
adult lung	GIBCO	ALG001	83-84 553 598 642-644 650 747 975 1004 1009 1015-1022 1449 1516 1816 2161 2186 2215 2343 2444 2968 3056 3158 3160-3162 3345 3466 3503 3519 3566 3568-3573 3595 3709 3797 3810 3839 3885 4025 4039 4059 4194 4405 4622 4755 4767 4770 4797 4865 4940 4949 4963 4972 4987 4991 4998 5001 5117-5129 5171-5172 5233 5401 5481 5581 5724 5788 5938 5960 6123 6125 6140 6215 6322 6336 6343-6348 6371 6886 6966-6967 7024 7143 7275 7441 7444 7502 7522-7523 7749-7752 8093 8110 8140 8336 8345 8375 8378 8449 8535 8612 8622 8635 8648 8651 8674 8702 8707-8708 8838-8839 8898 8941-8942 8944 8948 8951 9076 9128 9457 9466 9475 9497 9503 9544 9567 9703 9756 9758 9808-9810 9843 9873 9926 9999 10161 10175 10275 10329 10344 10400 10458 10554-10555 10796 10998 11032 11046 11069 11132 11144-11145 11483 11627 11818 11837 11901 12006 12028 12039 12041 12087-12092 12146 12363 12375 12424 12438 12539 12570 12601 12617 12661 12670 12674 12688 12738 12749 12754 12904 12940 12974 13062 13173 13195 13232 13234 13260 13295 13307 13456 13472 13477 13490 13494 13497-13498 13569 13584 13592 13660 13663-13666 13670 13743 13776 13882 13885 13889 13891
lymph node	Clontech	ALN001	83 142 364 487 495-497 554 629 645-646 648-650 716 938-951 953-962 1134 1516 1549 1671 1774 1976 2138 2225-2232 2234-2235 2340 2879 3019-3020 3057 3103 3163-3165 3272 3356 3498-3505 3507-3513 3515-3516 3538 3548 3628 3697 4194 4201 4253 4405 4641 4687-4688 4771-4772 4783 4808 4845 4963 4972 4987 4998 5042-5065 5076 5163 5504 5523 5835 5895 5917 6027 6142 6192 6272 6288 6308-6311 6313-6314 6765 6805 6871 6949-6951 7700-7705 7707-7709 8085 8088 8091 8110 8235 8375 8387 8432 8629 8631 8633 8648 8677 8709 8713-8715 8830 8863 8887-8889 8891-8896 8943-8944 8966 9010 9076 9111 9128 9142 9222 9455 9472 9520 9544 9734 9774 9780 9811 9905-9907 9928 9939 9999 10027 10129 10296 10439 10452 10501 10543-10545 10679 10777-10778 10788 10891 10999-11000 11111-11113 11115 11130 11145 11344 11406 11513 11584 11885 12006 12028 12050-12056 12143 12256 12363 12405 12442 12570 12674 12690 12913 12917 12940 12968-12970 13173 13220-13221 13394 13400 13492 13503 13511 13533 13630 13642 13645 13713 13868 13885 13889 13891
young liver	GIBCO	ALV001	211 498 598 651-653 1008-1014 1193 1264 1575 1976-1977 2131 2161 2254-2255 2269 2363 2568 2617 2627 2633 2636 2961 3021 3059 3093 3166 3215 3272 3356 3378 3426 3479 3559-3565 3590 3597 3627 3634 3673 3709 3797 3810 3885 3993 4006 4011 4019 4025 4194 4246 4253 4277 4422 4426 4431 4434 4437-4438 4634 4654 4687 4714 4722 4755 4773-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			4774 4808 4957 5001 5037 5052 5113-5115 5163 5167 5171 5508 5510 5526 5581 5616 5693 5726 5757 5772 5788 5928 6125 6141-6142 6193 6288 6340-6342 6589 6765 6815 6886 6901 7024 7118 7141 7442 7524 7735 7747-7748 7792 7852 7870 7924 7930 8066 8110 8120 8262 8375 8378 8518 8520 8648 8652 8691 8702 8716-8718 8816 8838-8839 8938-8940 9097 9119 9263-9264 9301 9317 9424 9454-9455 9544 9689 9703 9734 9758 9775 9925 9947 10393 10477 10769 10774 11001 11132 11140-11142 11173 11208 11318 11406 11587-11588 11725 11804 11902 12081-12086 12118 12160 12178 12375 12377 12405 12424 12479 12511 12570 12599 12601 12618 12670 12674 12688 12737 12830 12913 12916 12933 12953 12977 12993 13047 13062 13073 13126 13169 13174 13230-13231 13295 13488-13489 13494 13498-13499 13502 13506-13507 13575 13592 13646 13660-13662 13670 13866 13868-13869 13882 13888 13901
adult liver	Invitrogen	ALV002	6 25 60 142 598 1004 1213 1296-1301 1381 1493 1513-1514 1977 2139 2161 2269 2310 2335 2342 2400 2413 2449 2458 2497 2568-2569 2576 2636 2809 2827 2880 2926 3471 3484 3503 3597 3722 3792-3796 3885 3994-3995 4025 4060 4201 4326 4422 4426 4618 4689 4767 4786 4985 4998 5093 5163 5337-5339 5511 5526 5581 5693 5701 5724 5726 5757 5895 5922-5923 5979 6020 6027 6083 6125 6192 6195 6253 6333 6346 6483-6485 6716 6765 6797 6975 7169 7422 7468 7614 7642 7807 7932 8089 8110 8140 8262 8323 8378 8633 8677 8711 8740 8759 8786 8834 8839 8944 9002 9128-9129 9264-9265 9282 9466 9484-9485 9517 9677 9700 9773 10007 10046-10047 10135 10616 10669 10777 11032 11114 11132 11142 11194 11271 11389 11462 11502 11560 11587 11602 11818 11980 11989 12086 12160 12195 12253-12254 12316 12327 12363 12378-12379 12411 12424 12511 12570 12630 12693 12706 12788-12789 12840 12863 12913-12914 13047 13062 13072-13073 13090 13242 13256 13378-13379 13389 13479 13487 13489 13497 13555 13734-13738 13859 13864 13889
adult liver	Clontech	ALV003	346 2529 3548 3797 3885 4438 4940 5101 5801 5902 9597 10769 10778 11587 11927 12086
adult ovary	Invitrogen	AOV001	6 13 40 67 83 88 142 196 444 522 555-560 577 579 598 609 654-657 659-665 667-668 708 740 745-749 751-754 756-761 806 841 871 1004 1023 1028-1034 1036-1039 1041-1047 1055 1087-1115 1119 1151-1162 1298 1346 1359 1456 1493 1507 1516 1658 1697 1752 1774 1812 1826 1848 1914 1974 1976-1977 1979 1983 2099 2111 2138 2153-2155 2161 2175 2180-2181 2186 2258-2267 2284-2286 2288-2290 2298-2299 2342-2343 2449 2483 2523 2529 2555 2569 2576 2591 2597 2618 2708 2750 2783 2818 2839 2926 2961-2962 2988 3007 3022 3031 3060-3061 3063 3069 3166-3172 3174-3175 3177-3189 3226 3272 3286 3288-3289 3291-3293 3295-3299 3435 3450 3460 3479 3486 3503 3535 3577-3589 3612 3627 3629 3631-3643 3673 3675 3720 3722 3737 3797 3806 3810 3812-3813 3839 3885 3985 4006 4012-4014 4023 4025 4040 4043 4060 4078 4133 4192 4194 4201 4246 4269 4277 4282 4341 4362 4381 4383 4385 4388 4390-4391 4396 4399 4402 4405 4410 4434 4465 4473 4515 4543 4582 4610 4654 4689 4694 4715-4717 4722 4759 4767 4775-4782 4836-4837 4873-4878 4880 4940 4944 4957 4960 4962-4964 4972 4975 4981 4985 4998 5001-5002 5018 5030 5037 5040 5045 5075-5076 5093 5108 5135-5136 5138-5146 5148 5163 5191-5209 5213 5238-5245 5262 5374 5380 5451 5455 5490 5503 5505 5521 5523 5527 5560 5581 5674 5701 5724 5757 5778 5783 5788 5794 5817 5874 5894-5895 5902 5906 5908-5909 5916-5917 5923 5932 5938 5979 5989-5990 6005 6027 6048 6054 6057 6061 6078

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			6082 6086 6125 6140 6143-6144 6165 6171 6192 6194-6200 6215 6235-6239 6241 6270 6288 6307 6333 6336 6354-6360 6362-6364 6374 6385-6392 6410 6412 6415-6417 6458 6461 6508 6555 6583 6595 6598 6604 6607 6624 6626 6629 6708 6765 6767 6793 6797 6799 6801 6805 6825 6860 6870-6871 6873 6886 6888 6900 6902-6904 6915-6917 6947 6969-6970 6975 6980-6987 6998 7054 7058 7104 7109 7118 7133 7137 7141 7155 7166 7169 7171 7191 7204 7281 7386 7408 7431 7443 7468 7471-7472 7494 7507 7525-7528 7604-7609 7614 7642 7687 7716 7729 7762-7763 7765-7767 7769-7771 7773-7775 7788 7802-7818 7829-7835 7838-7841 7864 7905 7937 7977 8007 8035 8077 8088 8093 8095 8110 8120 8134 8139-8140 8143 8156 8235 8246 8262 8292 8320 8336 8345 8355 8358 8363 8365 8368 8370 8372 8375 8387 8444 8452 8456 8486 8520 8593 8610 8633-8634 8653 8719-8724 8773-8777 8779 8781-8782 8792 8798 8813 8830 8839 8853 8863 8870 8876 8898 8928 8944 8950 8956-8961 8963-8971 8989 9004-9007 9009-9019 9042-9046 9055 9076 9097 9128 9134 9185 9264 9283-9284 9291 9313 9358 9424 9445 9454-9455 9457 9460 9466 9471-9472 9475 9480 9511 9517 9533 9539 9542 9544 9553 9565 9577 9586 9612 9626 9677 9694 9703 9722 9734 9757-9758 9773 9776-9778 9791 9812-9819 9827 9838 9843 9848-9853 9922 9927-9940 9942 9962-9970 9989-9992 9994-9997 10007 10027 10097 10149-10152 10249 10252 10275 10298 10333-10334 10346 10349 10351 10383 10386 10417 10438-10439 10452 10459 10470 10478-10480 10493-10498 10557-10562 10564-10565 10580-10586 10594-10596 10645 10658 10676 10679 10767 10772 10774 10778 10782 10788 10791 10954 10971-10972 11002-11004 11032 11047-11052 11061 11069 11083 11108 11130 11132 11151-11153 11155 11157-11158 11167 11170 11172-11176 11193-11196 11274 11343 11369 11406 11411 11431 11588 11596 11600 11604 11606-11607 11618 11620-11621 11629 11668 11814 11860 11865 11876 11903-11909 11932 11949-11954 11980 11985 12000 12006 12017 12022 12033 12039 12080 12083 12092-12093 12095-12104 12127-12129 12131-12142 12146 12159-12160 12162-12163 12178 12213 12216 12236 12256 12260 12305 12327 12363 12368 12371 12375 12379 12387 12401-12402 12405 12424 12430 12467 12520 12522 12546 12570 12576 12590 12594 12599 12605 12609 12611 12615 12617-12618 12630 12636 12643 12657-12658 12663 12670 12674 12685 12688 12691 12693 12702 12705-12707 12713 12724 12729 12735 12740-12745 12749 12754-12760 12765-12770 12777 12842 12848 12875 12904 12906 12910 12913 12916-12917 12935 12940 12955 12957 12963 12972 12977-12980 12982-12984 12992-12994 12997-13000 13020 13034 13047 13051 13054 13056 13060 13062 13066 13070 13072-13073 13075 13082 13090 13092-13093 13104 13126 13136 13141 13175-13177 13179-13180 13193-13194 13196-13197 13202 13236-13241 13243 13245 13249 13254 13263-13275 13277 13286-13292 13295 13319 13351 13377 13389-13391 13394 13420 13436 13456 13477 13479 13488 13492 13494-13495 13497-13499 13502-13503 13506 13512 13516 13533 13546 13549 13554-13555 13575 13590 13597 13600-13601 13613 13616-13620 13627 13631 13644-13645 13649 13659-13660 13670 13674-13679 13693-13699 13707-13709 13713 13782 13803 13864 13866 13868-13869 13872 13875 13882-13885 13888-13889 13892
adult placenta	Clontech	APL001	669-671 1006 1134 1184 1551 2053 2090 2156-2158 2342 2490 2716 3001 3064-3065 3190 3272 3625 3670 4019 4055 4194 4201 4246 4641 4718 4720-4721 4783-4786 4957 5523

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			5536 5581 5788 5895 6145-6146 6201-6202 6358 6803 7049 7204 7529 8367 8375 8609 8611 8630 8725-8727 9097 9480 9734 9820 10319 10460 10579 10795 10895 11519 11723 11737 11776 11827-11828 11959 12403 12546 12570 12663 12848 12875 12913 13051 13394 13782 13869 13889
placenta	Invitrogen	APL002	142 561 1023 1551 1976 2449 2569 2614 3191 3340 3592 3668 4060 4346 4767 4787-4789 4892 4985 5536 5674 5693 5757 5772 5794 5906 6147 6203 6215 6598 6788 6799 6979 7158 7530-7531 8139 8361 8611 8863 9457 9484 9517 9563 9703 9758 9821 9927 9939 9942 10481 10557 10778 10802 10973 11006-11011 11723 11731 11737 11776 11807 11866 11910-11916 11959 12379 12403 12590 12685 13020 13054 13392 13489 13533 13554-13555 13867 13882 13888 13891
adult spleen	GIBCO	ASP001	463 499 562-563 598 672 990-993 995 1004 1020 1346 1515-1516 1556 1774 1877 1914 1977 2159 2161 2215 2248 2257 2363 2529 2569 2614 2617 2812 2827 3093 3192 3195 3199 3272 3286 3350 3356 3434 3470 3503 3537 3544-3546 3548 3604 3634 3689 3709 3720 3722 3796-3798 3812 3839 3885 4006 4011 4022 4025 4043 4194 4201 4246 4253 4277 4388 4391 4396 4405 4434 4641 4689-4690 4781 4957 4987 4998 5001 5017 5030 5037 5052 5076 5118 5512-5513 5523 5526 5701 5740 5778 5788 5796 5801 5895 5938 6005 6064 6125 6140 6147 6192 6204 6272 6329 6458 6551 6590 6607 6873 6886 6931 6958 7018 7058 7109 7118 7133 7171 7233 7506 7532-7533 7561 7586 7733 7736 7807 7842 7845 7933 8065 8085 8093 8095 8110 8116 8139-8140 8226 8235 8262 8323 8359 8363 8365 8368 8372 8456 8535 8648 8702 8792 8816 8838-8839 8858 8863 8865 8876 8923-8928 8943 8950-8951 9128 9257 9264 9317 9455 9457 9466 9474 9544 9560 9773 9781 9919-9920 9927 9939 9942 10000 10048-10049 10155 10175 10267 10275 10296 10331 10549 10566 10774 10777 10828 10944 11032 11061 11108 11113 11132 11173 11272 11401 11406 11620-11621 11804 11818 11861 11917 12006 12039 12041 12073-12074 12078 12118 12127 12236 12327 12375 12387 12402 12405 12424 12426 12511 12546 12560 12570 12601 12617-12618 12670 12674 12734-12735 12749 12772 12790 12913 12916-12917 12972 12977 13002 13021 13066 13082 13084-13085 13169 13228-13229 13256 13260 13295 13316-13317 13394 13456 13488-13489 13494 13497 13512 13516 13546 13645 13649 13651 13659 13679 13739 13776 13864 13866-13867 13888
testis	GIBCO	ATS001	50 142 500-502 564-565 598 673-678 963-968 1193 1346 1556 1671 1826 1968 1977 2160-2161 2236-2237 2555 2590 2597 3024 3093 3196-3202 3284 3378 3517-3522 3524 3720 3797 3839 3885 4006 4025 4055 4109 4253 4354 4377 4388 4405 4434 4473 4480 4543 4688 4792 4940 4957 4962 4964 4975 5001 5018 5030 5037 5067-5072 5076 5118 5171 5481 5525 5535 5740 5757 5788 5907 5924 6125 6147 6257 6316-6322 6324 6607 6886 6952 7058 7141 7288 7445 7484 7534-7540 7677 7710-7711 7713-7717 7735 7792 7815 8087 8110 8120 8262 8368 8370 8375 8535 8577 8648 8655 8702 8729 8838-8839 8848 8898-8906 8936 8939 8966 9076 9087 9244 9264 9304 9455 9466 9485 9542 9567 9703 9758 9822-9825 9908-9910 9928 10462 10475 10478 10482 10551 10583 10679 10772 10778 10795 10955 10974 11108 11113 11116 11118-11120 11132 11173 11239 11425 11606 11620 11629 11695 11807 11862 11918-11919 12000 12006 12017 12033 12057-12061 12375 12405 12424 12570 12577 12599 12601 12605 12609 12668 12692 12730-12731 12748 12904 12917 12923 12929 12935 12940 12956 12971 13047 13051 13073 13090 13169 13222-13225 13249 13394 13442 13456 13479 13494

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			13497 13502-13503 13506 13512 13515 13533 13549 13555 13575 13613 13616 13627 13644-13646 13660 13866 13875 13882 13889
Genomic DNA from BAC 63118	Research Genetics (CITB BAC Library)	BAC001	8711
adult bladder	Invitrogen	BLD001	731 1710 1779-1791 2076 2367 2479-2481 4067 4208-4214 4681 4767 4775 5163 5169 5553 5560 5581 5674 5684-5688 5794 6082 6244 6716 6718-6721 6799 6931 7045 7094-7097 7967 8110 8226 8249-8258 8364 8648 8737 9012 9097 9425-9428 9430 9626 9703 9928 10007 10261 10557 10734-10735 11459 11504-11506 12000 12212 12512-12513 12515-12517 12670 12735 12914 13085 13439 13512 13687 13838-13841 13866
bone marrow	Clontech	BMD001	11 70 83 85 142 150 162-184 186-198 200-210 230-243 245-277 279-281 306 370 373-387 389 433-435 438-440 457 483 503-510 566 568-569 595-596 598 609 616 679-683 740 770 843-857 859 861 864 867-876 880-884 886-890 892-893 896 900 902-920 1000 1004 1116-1120 1122-1126 1128-1134 1184 1211 1346 1359 1516 1552-1554 1575 1583 1671 1724 1774 1877 1894 1927 1974 1976 1983 2012-2024 2031 2033-2038 2040 2043-2044 2084-2088 2111 2120-2121 2132-2133 2137-2138 2161 2163-2164 2186 2189 2199-2200 2202-2203 2205-2210 2213-2218 2266 2291-2295 2342-2343 2414-2416 2444 2529 2555 2566 2569 2575-2576 2591 2597 2652 2681-2709 2711-2716 2735-2738 2740-2744 2746-2748 2750-2756 2758-2759 2761-2764 2766-2768 2770 2772-2781 2783-2787 2806 2812-2813 2816 2873 2875 2880-2881 2883-2885 2887-2897 2899 2901 2926 2948 2954 2958 2977 2984 3026-3029 3032 3068-3069 3071 3081 3093 3096-3097 3139 3203-3205 3207 3238 3257 3272 3282-3283 3286 3317 3345 3356 3404-3418 3422-3434 3437-3438 3440 3442 3447 3449-3450 3456-3459 3461-3464 3466-3473 3483 3497 3535 3538 3548 3557 3572 3588 3593 3600 3604-3605 3610 3612 3625-3627 3634 3644-3647 3649-3651 3653-3657 3709 3711 3722 3725 3727 3737 3797 3804 3808 3810 3839 3899 4005-4006 4011 4015-4019 4023 4025 4040 4043 4058 4060 4129 4132 4192 4194 4201 4246 4253 4277 4282 4367 4403 4466-4472 4474-4477 4479-4484 4486-4490 4492-4497 4509 4512-4540 4582 4595 4597-4606 4608-4615 4622 4642 4648 4650 4654 4667 4691-4695 4723 4741-4742 4759 4767 4783 4794-4796 4808 4836 4928-4929 4931-4934 4938 4940 4942 4944-4948 4950-4953 4955-4960 4962-4964 4967-4971 4973 4976-4979 4985-4987 4992-4998 5000 5004-5016 5030 5037 5052 5075-5076 5093 5143 5149-5151 5163 5169 5210-5215 5217-5226 5262 5317 5357 5503 5523 5535 5537-5540 5560 5604 5695 5740 5748 5766 5788 5796 5801 5862 5874 5895 5906 5908 5938 5965-5967 5969-5979 5991-5994 5996-6005 6007-6009 6027 6047 6057 6065 6067-6068 6071-6072 6082 6086 6089-6091 6112 6125-6127 6148-6152 6195 6214 6233 6257 6269-6276 6278-6280 6282-6283 6285-6289 6292-6300 6321 6374 6393-6403 6508 6555 6605 6607 6722 6730 6788 6815-6822 6826-6832 6836 6859-6861 6867 6870-6871 6873-6874 6905 6915 6929 6933-6937 6940-6942 6944-6945 6988-6992 7051 7126 7155 7166 7169 7199-7210 7212-7219 7226 7233-7253 7256-7264 7275 7317-7318 7320 7322-7326 7347 7349 7373-7374 7379 7382 7446-7448 7473-7474 7491 7502 7542 7544 7569 7577 7579 7663-7671 7674 7677-7678 7680-7684 7686 7689-7696 7742 7768 7791 7815 7819-7824 7847 7864 7930 7995 8065 8085

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			8088 8090-8093 8095 8110 8120 8134 8139 8224 8226 8235 8241 8246 8262 8336 8354 8359 8363 8365 8368 8375 8421-8427 8429-8440 8444 8452 8456-8460 8463-8465 8467-8476 8478-8489 8538-8539 8542-8551 8563 8566 8577 8593-8595 8597 8603 8633 8635-8637 8648 8656-8657 8671 8677 8679 8701 8792 8796 8817 8830 8840-8845 8847-8849 8852 8855-8860 8863-8868 8870 8873 8875 8879-8881 8898 8935 8944 8963 8965 8995 9010 9012 9020-9031 9049 9051 9066 9076 9092 9097-9098 9128 9257 9264 9285-9288 9304 9317 9339 9381 9399 9401 9454-9455 9466 9471-9472 9474-9476 9516-9517 9519 9529 9544 9622 9630-9643 9646 9650-9665 9689 9705-9712 9730 9734 9740 9758 9779-9781 9826 9885-9889 9892-9902 9928-9929 9939 9942 9972-9976 9978-9982 9999-10000 10068 10153 10175 10249 10275 10277 10284 10287 10296 10319 10321 10343-10344 10346 10364-10374 10379-10382 10385-10394 10414 10416 10426 10429 10439 10447 10452 10467-10468 10475 10483 10520 10522-10526 10528-10529 10531-10535 10587-10590 10595 10616 10677-10679 10691 10750 10760 10767 10772 10778 10788 10794-10795 10828 10831-10837 10846-10847 10849-10851 10853-10858 10891 10895 10897 10899-10903 10905-10907 10931-10933 10956-10959 10975-10976 11012 11032 11046 11057 11059 11061 11077-11082 11085-11086 11088-11098 11108 11113 11132 11145 11177-11181 11194 11208 11289 11344 11401-11402 11513 11618 11620 11631 11673-11675 11677-11690 11695 11697-11704 11706 11708-11715 11736 11742 11762-11772 11774 11795-11796 11802-11804 11817-11818 11829 11839-11842 11863 11920 11946 11992-12002 12004-12008 12010 12012-12013 12015-12019 12022-12029 12033 12041 12081 12142-12149 12160 12178 12195 12200 12231 12236 12283 12305 12308 12316 12327 12363 12368 12404-12405 12424 12426 12430 12439 12527 12546 12570 12576 12590 12608 12610 12616 12634-12636 12641-12645 12654-12655 12657-12659 12669-12670 12674 12721-12724 12754 12761 12834 12840 12842 12904-12905 12910-12911 12913 12916-12917 12922-12923 12925-12928 12933 12938 12956 12966 12977 12989-12990 12993 12995 13020 13033 13047 13060 13062 13064 13066 13095 13098 13103-13105 13107 13111-13113 13129 13135 13137 13141 13150 13163 13173 13181 13213-13217 13234 13276-13278 13280 13391 13394 13400 13456 13472 13477 13490 13492 13494 13498 13501-13502 13506-13507 13511 13516 13528 13530-13534 13538 13540 13546 13551-13552 13554-13555 13560 13585-13586 13594 13613 13630 13633-13639 13644-13646 13649 13659-13660 13670 13673 13679 13713 13775 13795 13866 13872 13875 13882-13885 13889 13891 13893
bone marrow	Clontech	BMD002	51 242 442 654 1004 1134 1841-1904 1908 1927 2023 2107 2215 2342 2408 2507-2529 2576 2597 2806 2866 3286 3434 3722 3736-3737 3817 3823 3839 4060 4246 4258-4290 4389 4396 4411 4618 4641 4828 4836 4957 4987 5030 5037 5052 5108 5163 5526 5711 5743-5772 5774-5804 5813 5895 5939 6131 6236 6266 6287 6306-6307 6333 6655 6675 6728 6730 6750-6762 6764-6769 6788 7049 7106-7117 7137 7156 7738 8068 8085 8134 8156 8290-8297 8299-8305 8307-8312 8323 8368 8482 8499 8858 8863 8865 8928 8935 9076 9111 9128 9465-9502 9626 9700 9703 9923 9928-9929 9942 10277-10280 10282-10284 10286-10297 10396 10434 10515 10551 10645 10675 10691 10744-10753 10772 10778 11057 11098 11108 11132 11232 11252 11519 11536-11553 11606 11620 12033 12039 12146 12260 12305 12387 12402 12405 12500 12554-12578 12594 12599 12608 12674 12754 12777 12839 12895-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			12904 12916 12923 13020-13021 13043 13057-13064 13169 13249 13446-13454 13479 13492 13494 13513 13533 13555 13659 13713 13775 13803 13863-13875 13877-13887 13889 13891
bone marrow	Clontech	BMD004	2249 2529 3286 3494 3548 3551 3797 3839 4025 4058 4201 4277 4282 5052 5108 6545 6961 8262 8898 9474 10000 11098 11818 13021 13893
bone marrow	Clontech	BMD007	8539 9780 9927 13021
adult colon	Invitrogen	CLN001	319 346 487 731 799 1792 1848 2050 2161 2449 2482-2483 3431 3901 4215-4217 4940 4957 4987 5163 5239 5560 5689-5695 5865 5911 5923 6722 6765 7098 7815 7864 7880 8110 8259-8262 8486 8597 8951 9484 9529 9542 9556 10376 11507-11508 11617 11869 12127 12236 12424 12518-12523 12601 12610 12777 12976 13062 13073 13367 13440 13507 13512 13630 13713 13843-13844 13864 13868-13869 13888
Mixture of 16 tissues – mRNAs*	Various Vendors*	CTL016	6815 10776 12977 13064 13512
Mixture of 16 tissues – mRNAs*	Various Vendors*	CTL021	1671 6738 8432 8648 8863 8944 9511 10769 13021 13062 13064
adult cervix	BioChain	CVX001	50 67 142 158 308 332 346 475 598 654 895 1004 1086 1286 1449 1516 1671 1698 1701 1711-1756 1758-1776 1828 1848 1959 2134 2186 2257 2267 2343 2408 2414 2468-2474 2476-2478 2608 2716 3002 3136 3166 3191 3199 3529 3535 3554 3572 3627 3722 3737 3777 3797 3839 3985 4158-4176 4178-4195 4197 4199-4207 4246 4277 4391 4396 4434 4641 4667 4759 4783 4828 4885 4940 4957 4963 4987 4998 5001 5038 5075 5108 5163 5293-5294 5455 5481 5523 5552 5581 5646-5652 5654-5659 5661-5671 5673-5681 5687 5701 5711 5723 5740 5788 5794 5848 5902 5908 5923-5924 5964 6020 6052 6057 6062 6091 6106 6112 6125 6129 6181 6350 6371 6374 6410 6446 6458 6504 6508 6512 6551 6598 6686-6687 6689-6705 6707-6715 6788 6873 6893 6917 6998 7008 7045 7078 7084-7093 7095 7130 7141 7148 7169 7204 7507 7579 7608 7675 7733 7768 7815 7871 7880 7893 8078 8138 8209-8215 8217-8236 8238-8242 8244-8248 8298 8345 8370 8444 8456 8486 8499 8535 8558 8592 8633 8635 8648 8669 8679 8742 8853 8863 8870 8898 8921 8939 8948 9012 9061 9098 9107 9128 9137 9153 9304 9308 9317-9318 9355 9385-9391 9393-9403 9405-9406 9408-9418 9420-9422 9457 9466 9475 9510 9539 9612 9734 9773 9927-9928 9939 9947 9960 10110 10175 10230-10256 10258-10259 10267 10274 10319 10329 10344 10491 10496 10540 10616 10660 10691 10722-10732 10778 10782 11055 11145 11217 11376 11462 11477-11489 11491-11503 11519 11584 11604 11695 11853 11869 11891 11980 12006 12066 12081 12127 12160 12195 12216 12240 12266 12308 12363 12379 12402 12405 12424 12438 12483-12494 12496-12510 12579 12605 12610-12611 12617-12618 12643 12653 12670 12674 12688 12691 12703 12707 12735 12740 12754 12830 12840 12866 12870-12881 12883-12884 12905

* The 16 tissue-mRNAs and their vendor source, are as follows: 1) Normal adult brain mRNA (Invitrogen), 2) normal adult kidney mRNA (Invitrogen), 3) normal adult liver mRNA (Invitrogen), 4) normal fetal brain mRNA (Invitrogen), 5) normal fetal kidney mRNA (Invitrogen), 6) normal fetal liver mRNA (Invitrogen), 7) normal fetal skin mRNA (Invitrogen), 8) human adrenal gland mRNA (Clontech), 9) human bone marrow mRNA (Clontech), 10) human leukemia lymphablastic mRNA (Clontech), 11) human thymus mRNA (Clontech), 12) human lymph node mRNA (Clontech), 13) human spinal cord mRNA (Clontech), 14) human thyroid mRNA (Clontech), 15) human esophagus mRNA (BioChain), 16) human conceptional umbilical cord mRNA (BioChain).

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			12913-12914 12917 12923 12951 12955 12957 12965 12989-12990 12993 13002 13020 13035 13045-13050 13062 13064 13072 13092 13136 13141 13174 13197 13254 13283 13307 13391 13428 13430-13437 13442 13473 13479 13492 13494-13495 13497-13498 13502 13532 13554-13555 13575 13590 13597 13613 13616 13627 13644 13679 13713 13775 13829-13837 13866 13868-13869 13872 13884 13888 13891
diaphragm	BioChain	DIA002	731 1346 3548 3711 3885 4282 4654 5895 6873 8120 8931 8936 9455 11132 11818 12405 12609
endothelial cells	Strategene	EDT001	21 51 67 83 332 569 598 609 762 796 1004 1024-1026 1086 1561 1848 1928 1959 1976-1977 1983 2138 2161 2166 2257 2282-2283 2417 2483 2490 2555 2569 2614 2926 3042 3189 3191 3272 3300-3303 3426 3494 3503 3548 3574-3576 3605 3627-3628 3673 3709 3720 3722 3737 3797 3839 3885 4005 4011 4019 4055 4133 4192 4246 4269 4282 4340 4354 4365 4384 4388 4399 4405 4410-4411 4434 4543 4641 4654 4767 4797-4799 4802 4881-4882 4885-4886 4888 4940 4957 4964 4972 4985 4998 5002 5017 5030 5076 5103-5104 5132-5133 5163 5167 5183-5185 5187-5190 5380 5523 5527 5535 5541-5542 5544 5674 5684 5693 5724 5766 5778 5788 5794 5796 5874 5895 5916 5923 5928 5938 6005 6048 6057 6068 6082 6165 6205 6215 6240-6241 6307 6321-6322 6349-6353 6383 6458 6595 6598 6606-6607 6765 6799 6805 6815 6860 6871 6873 6890 6918 6968 6972 6976 6979-6980 6998 7058 7067 7104 7113 7116 7137 7139 7169 7275 7468 7613 7716 7755-7757 7759-7760 7797 7799-7801 7930 8077 8084 8093-8095 8120 8139 8235 8262 8320 8323 8335-8336 8345 8354 8358 8363-8364 8370 8372 8375 8387 8452 8592 8648 8786 8788 8792 8813 8863 8898 8944-8955 8965 8996-9001 9051 9076 9097 9128 9264 9289-9291 9304 9315 9414 9455-9456 9466 9472-9473 9475 9484 9504 9517 9529 9542 9563 9570 9626 9703 9780-9781 9843 9927 9939 9961 10000 10027 10154 10267 10285 10321 10330-10331 10342 10344 10349 10496 10500 10550 10556-10557 10579 10679 10772 10776 10778 10788 10795 10802 11013 11132 11136 11146-11149 11406 11483 11565 11588 11600 11606-11607 11615 11626 11807 11818 11932 11955 12006 12034 12037 12041 12044 12078 12092-12094 12123-12126 12150 12213 12375 12381 12387 12405 12411 12424 12426 12522 12570 12576 12590 12601 12610 12612 12615 12617-12618 12663 12670 12674 12707 12729 12739 12749 12753-12754 12777 12830 12842 12913-12914 12916-12918 12929 12940 12972 12975-12977 13002 13024 13047 13051 13054 13062 13064 13082-13084 13090 13092 13094 13123 13126 13136 13195 13235 13263 13380 13389 13392 13394 13400 13456 13479 13488-13489 13492 13494-13499 13502 13506-13507 13514 13516 13546 13555 13568 13575 13590 13592 13613 13616 13621 13630 13649 13659-13660 13667-13668 13670-13673 13678-13679 13690-13692 13713 13796-13797 13838 13866-13869 13872 13882 13884 13888 13893
Genomic clones from the short arm of chromosome 8	Genomic DNA from Genetic Research	EPM001	150 2023 2327 2490 4109 4783 5503 5560 10267 10760 12017 12160 12557 12582 12923 13020 13514
Genomic clones from the short arm of chromosome	Genomic DNA from Genetic Research	EPM003	5560 12017 12146

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
8			
Genomic clones from the short arm of chromosome 8	Genomic DNA from Genetic Research	EPM004	4783 4798 5560 10817 11926 12017 12160
esophagus	BioChain	ESO002	999-1000 2449 3272 3315 3548 3550 3634 3697 3796 4011 4025 4058 4201 4282 5106 5163 5553 6082 6873 7739 9304 10296 11133 11818 12033 12570 13869
fetal brain	Clontech	FBR001	51 142 1184 3664 4060 4109 4940 5021 5270 5523 5553 6112 6805 6908 7294 8558 9457 10376 11059 11985 12006 12122 12160 12754 13438 13507 13888
fetal brain	Clontech	FBR004	60 2704 3711 4025 4109 4783 5001 6082 7597 9010 9504 9949 11837 12033 12039 12363 12705 12905 13020 13503 13512 13891
fetal brain	Clontech	FBR006	6 60 67 598 800 932 1004 1170 1793-1794 1796-1797 1799-1805 1905-1914 1916-1958 1974 1976 1979 1983 2057 2129 2174 2221 2407 2444 2449 2484-2492 2530-2554 2556-2561 2563 2576 2857 3064 3207 3479 3556 3673 3709 3722 4060 4078 4157 4218-4221 4223-4224 4277 4291-4334 4338 4355 4364 4369 4431 4957 5001 5109 5270 5380 5553 5634 5696-5706 5711 5724 5766 5788 5794 5801 5805-5832 5834-5879 5882-5901 5936 5990 6057 6723-6732 6765 6770-6791 6797 6805 6894 7049-7050 7100-7102 7105 7118-7123 7125 7127 7169 7905 8263-8265 8267-8273 8294 8312-8333 8359 8361 8375 8452 8633 8664 8740 8757 8884 9010 9111 9432-9436 9503-9516 9518-9545 9547-9551 9556 9570 9577 9780 9895 9923-9924 9928 9942 10007 10027 10202 10263-10268 10276 10284 10298-10310 10329 10331 10496 10542 10595 10621 10736-10737 10755-10761 10772 10774 10795 11108 11132 11406 11483 11509-11523 11555-11582 11589-11590 11600 11606 11621 11713 11729 11807 11837 12006 12039 12044 12092 12113 12218 12231 12236 12327 12363 12398 12405 12465 12511 12524-12530 12576-12577 12579-12601 12729 12735 12754 12863 12869 12889 12906-12910 12914 12954 12973 13020-13021 13051-13052 13054 13065 13082-13083 13427 13445 13455-13470 13488 13490 13496 13498-13501 13507 13516 13560 13613 13630 13649 13708 13713 13769 13831 13845-13855 13868 13872 13882 13884 13888-13894 13896-13900
fetal brain	Clontech	FBRs03	1005 4405 5111 6337 6964 7742 13084 13864 13891
fetal brain	Invitrogen	FBT002	51 83 142 321 430 746 932 1054 1058-1065 1493 1833 1947 2273-2275 2299 2444 2449 2926 3479 3492 3885 4347 4354 4391 4405 4410 4434 4530 4804 4985 4998 5075 5160-5169 5380 5428 5466 5750 5788 5801 5895 6132 6215 6371-6374 6458 6598 6973-6974 7067 7096 7776-7778 7780-7782 7937 8143 8323 8361 8364 8372 8377 8452 8633 8977-8984 8986 9010 9142 9264 9332 9457 9474 9503 9511 9517 9539 9582 9827 9848 9927 9950-9953 10027 10161 10329 10430 10492 10573-10575 11014 11160-11164 11406 11628 11742 11814 11830 11985 12092 12112-12114 12116-12117 12127 12424 12511 12521 12570 12576 12643 12696 12735 12748-12751 12754 12830 12835 12913 12957 12977 12990 13002 13020 13062 13072 13083 13117 13254-13259 13377 13486 13489 13496 13499 13507 13590 13649 13685-13688 13713 13867 13888 13891 13893
fetal heart	Invitrogen	FHR001	1001 1004 2250 4025 6334 6765 7740 8933 8935 9457 9544 10000 11132 12599 12609 13021 13568 13656 13866
fetal kidney	Clontech	FKD001	142 346 364 511-517 570-572 574 598 685-690 969-970 972-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			981 983-987 1134 1346 2123 2167-2169 2238-2246 2342 2444 2483 2516 2555 2617 2728 2843 2876 3032 3049 3072 3206 3208 3282-3283 3525-3531 3533-3543 3548 3591 3709 3722 3797 3839 3878 4015 4019 4043 4246 4277 4367 4405 4696-4698 4725 4767 4805-4810 4940 4947 4957 4986 4998 5037 5056 5073-5080 5082-5091 5099-5100 5108 5258 5504 5523 5560 5923 6005 6207-6208 6225 6272 6288 6325-6332 6478 6603 6702 6793 6815 6906 6953-6959 7045 7058 7204 7355 7426 7449-7450 7520 7543-7546 7561 7587 7718-7732 7930 8077 8097 8262 8375 8387 8452 8520 8638 8658 8736-8737 8834 8863 8898 8907-8918 8922 8950 9010 9134 9257 9401 9457 9544 9597 9760 9781 9791 9828-9830 9912 9914-9918 10296 10440 10484 10546-10548 10772 11108 11121-11129 11131-11132 11170 11513 11638 11695 11923-11924 12006 12033 12062-12070 12072 12160 12405 12522 12570 12594 12599 12605 12626 12663 12670 12732-12733 12749 12848 12904 12914 12940-12941 12990 13020 13083 13188 13226-13227 13234 13263 13277 13280 13351 13391 13394 13491 13501 13512 13590 13644 13647-13650 13713 13782 13867-13868 13872 13875
fetal kidney	Clontech	FKD002	3286 5030 5037 5105 11108 12033 12490 12570 13494 13866
fetal kidney	Invitrogen	FKD007	3272 3806 4025 4253 4277 4654 5112 5535 5788 5801 8863 8935 9401 9466 10553 11628 11818 13494 13646 13866
fetal lung	Clontech	FLG001	79 2367 2395 3010 3460 3885 4828 4948 4962 5001 5723 5748 5902 5908 6186 6738 7051 7067 7677 7759 9264 9553 9700 10007 10478 11098 12017 12383 12417 12424 12749 12917 13020 13169 13472 13554 13644 13782 13835
fetal lung	Invitrogen	FLG003	142 319 364 629 1671 1806-1814 1816-1819 1877 2129 2161 2169 2367 2449 2493 2529 3191 3503 3610 4109 4225-4234 4367 4434 4957 5108 5380 5421 5581 5707-5710 5712 5714 5788 5801 6057 6733-6741 7034 7103 8274-8278 8365 8597 8948 9264 9327 9437-9442 9444 9466 9510 9525 9530 9539 9677 9773 9841 10007 10190 10198 10269-10271 10329 11519 11524-11527 11927 12531-12539 12848 12890 12904 13021 13072 13249 13445 13472 13489 13551 13575 13649 13670 13679 13856-13857
fetal lung	Clontech	FLG004	1003-1004 2597 5110 6963 9924 10552 11138-11139 12080 12990 13659
fetal liver-spleen	Columbia University	FLS001	-2 4-14 16-22 24 26 28-31 33-46 48-49 51-61 63-68 71-91 93-102 104-110 112-124 126-156 158-162 282-283 285-290 292-299 301-304 307-312 314-326 328-338 340-344 346-353 355-365 369 390-400 402 436 441 483 557 567 575-585 595 598 629 673 678 691-699 701-702 708 731 736 763-767 769-776 778-786 788-791 793-794 796 925 975 1004 1015 1023 1038 1068 1104 1134 1144 1184 1192 1216 1264 1298 1346 1482 1493 1516 1518-1521 1551 1556 1575 1583 1594 1636 1641 1707 1724 1774 1826-1829 1841 1858 1927 1959 1962-1965 1967-1972 1974-1979 1981-1998 2000-2009 2011 2045-2051 2053-2055 2057-2058 2060-2063 2065-2083 2089-2094 2100-2101 2161 2170 2174 2184-2194 2215 2222 2269 2290 2310 2342 2409 2411 2414 2444 2449 2458 2483 2490 2497-2498 2510 2516 2523 2529 2555 2562 2566-2576 2578-2586 2588-2591 2593-2601 2604 2607-2608 2611-2612 2614-2618 2620 2622-2642 2644 2646-2653 2655-2664 2666 2668 2670-2680 2696 2750 2788-2793 2795-2811 2814-2826 2828-2835 2837-2842 2844 2846-2848 2851-2858 2860 2862-2871 2876 2878 2893 2900-2905 2907-2909 2926 2929 2949 2952-2953 2959-2960 2984 2992 3032 3058 3069 3073-3076 3078-3080 3082 3093 3166 3194 3196 3207 3210-3211 3213-3214 3217-3225 3249 3257 3272 3282 3286-3287 3304-3307 3310-3311 3314-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			3317 3319 3321-3332 3334-3346 3356 3419 3426 3434 3446 3450 3455 3468 3470 3479 3484 3494 3503 3535 3537 3548 3552 3557 3572 3590 3594 3597 3604-3605 3610 3612 3625 3627 3634 3668 3670 3709 3711 3720 3722 3729 3737 3777 3797 3806 3808 3810 3813 3839 3885 3926 3990 3996-3997 4006 4009 4011 4019-4022 4025 4040 4043 4060 4078 4095 4109 4129 4192 4194 4201 4245-4246 4253 4261 4277 4282 4297 4335-4338 4340-4357 4359-4361 4364 4366-4367 4370-4372 4374-4377 4379-4382 4384-4386 4388-4392 4395-4396 4398-4414 4417-4424 4426-4455 4457-4458 4460-4465 4542-4543 4545 4547-4553 4555-4562 4565-4575 4577 4579 4581-4582 4585-4588 4593 4596 4607 4616-4629 4644-4645 4647 4654 4671 4676 4687 4689 4694 4721 4726-4729 4759 4767 4775 4783-4784 4788 4790-4791 4811-4819 4830 4837 4845 4862 4874 4889-4902 4930 4940 4948-4949 4957 4962-4965 4972 4985 4998 5022 5029-5030 5037 5040 5075-5076 5092 5108 5132 5152 5163 5167 5171-5172 5335 5380 5398 5473 5503 5514 5516 5523 5526-5527 5535-5536 5553 5581 5598 5604 5616 5674 5684 5691 5693 5711 5715 5724 5726 5748 5750 5757 5778 5788 5794 5801 5817 5865 5874 5894-5895 5902-5914 5916-5919 5921-5928 5930-5949 5951-5954 5956-5964 5966 5979 6005 6010 6013-6015 6017-6018 6020-6022 6024-6025 6028-6032 6034-6048 6052-6062 6064 6068 6073-6081 6086 6092-6094 6104 6112 6125 6135 6140 6147 6149 6153-6157 6165 6186 6195 6209-6211 6240-6241 6243-6248 6250-6253 6272 6276 6287-6288 6307 6313 6338 6371 6374 6419 6430 6446 6451 6458 6478 6496 6508 6545 6579 6595 6598 6607 6611 6624 6629 6642 6658 6695 6726 6728 6730 6738 6745 6754 6765 6767 6788 6793-6797 6799-6808 6810-6815 6833-6843 6845-6847 6850-6856 6858 6860 6862-6865 6870-6871 6875 6877 6891 6893-6894 6900 6907-6908 6917 6919-6925 6950 6979 7049 7058 7067 7077 7096 7109 7116 7118 7128-7131 7133-7135 7137-7141 7144-7147 7149-7150 7152-7153 7155-7179 7181-7192 7194-7198 7265-7271 7273-7280 7282-7305 7307-7312 7316 7327-7336 7375-7376 7383 7386 7451 7455 7468 7473 7475-7477 7479 7484 7548-7553 7555-7558 7561 7608 7617-7618 7620-7631 7642 7675 7687 7695 7716 7768 7809 7811 7839 7842 7864 7883 7933 7977 8069 8093 8095 8105 8110 8116 8139 8224 8226 8235 8241 8262 8323 8334-8340 8342-8350 8352-8355 8357-8359 8361-8408 8410-8411 8414-8419 8452 8456 8490-8498 8500-8502 8504-8526 8529-8530 8537 8540-8541 8553-8560 8563 8566 8568 8577 8592 8611 8639 8648 8659-8661 8664 8668-8669 8710 8738-8739 8741 8777 8792-8804 8830 8834 8839 8853 8858 8863 8865 8876 8898 8926 8935 8939 8950-8951 8957 8963 9010 9012 9076 9092 9097 9112 9119 9128 9257 9264 9266 9284 9291 9301 9304 9313 9318 9352 9375 9399 9414 9424 9445 9455 9457 9466 9473-9474 9476 9480 9484-9485 9511 9517 9528-9529 9536 9539 9542 9544 9552-9557 9559 9561-9564 9566-9572 9574-9577 9579-9630 9646 9655 9666-9667 9669-9671 9673 9675-9697 9700-9701 9703-9704 9713-9715 9718-9719 9722 9726-9729 9734 9752 9758 9761 9777 9780-9784 9813 9831 9854-9867 9869 9906 9927 9929 9939 9942-9943 9969 9983 10000 10007 10205 10267 10284 10296 10305 10311 10313-10316 10318-10325 10327 10329-10331 10333-10354 10356-10363 10395-10397 10400-10403 10405-10410 10413 10417-10422 10427 10446-10447 10463 10470 10485 10496 10502-10507 10515 10542 10557 10565-10566 10645 10679 10688 10691 10729 10750 10760 10762 10764-10769 10772-10774 10776-10782 10784 10786-10795 10797-10815 10817-10830 10859 10861-10869 10871-10879 10881-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			10893 10908-10916 10929 10934 10943 10948 10968 10971 10975 10977-10983 11005 11009 11011 11015-11018 11028 11031 11054-11063 11094 11108 11114 11132 11136 11142 11194 11208 11213 11246 11390-11391 11401 11406 11411 11414 11459 11462 11483 11508 11513 11521 11534 11559-11560 11565 11573 11585-11603 11606-11610 11612-11616 11618-11638 11640 11642-11657 11659-11667 11669-11670 11672 11699 11716-11718 11720-11722 11724-11741 11743-11758 11760-11761 11775-11787 11797 11799-11801 11818 11830 11862 11864-11866 11869 11876 11889 11911 11925-11928 11958-11961 11963-11976 11979-11980 12000 12006 12017 12019 12028 12033-12034 12041 12078 12086 12092 12122 12127 12159-12160 12178 12216 12218 12236 12245 12256 12260 12289 12303 12305 12308 12327 12330 12363 12375 12382 12387 12402-12405 12424 12426 12439 12442 12465 12470 12476 12511 12522-12523 12546 12551 12560 12570 12576 12590 12601-12602 12604-12619 12621-12633 12636 12639 12646-12653 12670-12671 12674 12682-12683 12688 12693 12696 12706 12709-12715 12735 12749-12750 12772 12790 12823 12830-12832 12835 12842 12849 12855 12863 12904-12905 12910-12911 12913-12916 12918-12924 12929 12931-12932 12934-12937 12939-12940 12954-12955 12965 12973 12976-12977 12989-12990 12993 12999 13020-13021 13035 13051 13054-13055 13060 13062 13066-13067 13069-13087 13089-13095 13097-13102 13114-13117 13119-13120 13122-13126 13134-13136 13141 13147 13156 13174 13193 13195-13196 13198-13201 13254 13260 13264 13277 13280 13295 13351 13368 13377 13391 13394 13400 13409 13412 13420 13456 13472-13477 13479-13484 13486-13492 13494-13496 13498-13499 13501-13510 13512-13516 13518-13527 13533 13541-13544 13546-13551 13553-13555 13560 13575 13587-13589 13597 13603-13604 13613 13616 13622-13623 13630-13631 13644-13646 13649 13659 13670 13679 13713 13743 13748 13769 13775 13782 13793 13803 13808 13818-13819 13858-13860 13864 13866-13869 13872 13882 13884 13888-13889 13891 13893 13901
fetal liver-spleen	Columbia University	FLS002	6 16 24 30-31 63 67 81 83 89 95 103 115 117 126 140 142 147 150-151 158 162 211 225 287 308 332 356 358 390-391 438-439 483 551 556 641 654 694 701 708 731 788 997 1006 1012 1047 1082 1151 1154 1178 1184 1208 1212 1356 1480 1507 1551 1556 1623-1647 1649-1667 1669-1671 1675 1677-1698 1828-1829 1858 1877 1885 1889 1914 1927 1947 1961 1997 1999 2028 2057 2070 2092 2098 2138-2139 2154 2161 2174-2175 2189 2191 2257 2282 2327 2342 2363 2399 2403 2409 2440-2441 2443-2450 2452-2459 2490 2498 2522-2523 2544 2555 2568 2575-2576 2592 2597 2605 2619 2623 2659 2806 2809 2812 2860 2869 2879 2903 2926 2932 2988 3031 3066 3075 3089 3188-3189 3286-3287 3319 3343-3344 3356 3426 3535 3548 3552 3554-3555 3583 3591 3610 3625 3634 3673 3709 3711 3720 3722 3839 3899 3926 3985 4005-4006 4011 4020 4025 4055 4058 4060 4078 4091-4116 4118-4126 4128-4142 4172 4194 4201 4253 4277 4282 4347 4350 4353-4355 4362 4367 4374 4379 4386 4391 4394 4396 4402 4426 4431 4435 4437-4439 4512 4553 4578 4586 4607 4622 4644 4654 4671 4758 4767 4783 4798 4836 4845 4899 4940 4948 4962-4963 4985 4991 4998 5001 5037 5108 5167 5171 5177 5198 5237 5293-5294 5380 5400 5523 5535-5536 5581-5591 5593-5599 5601-5613 5615-5624 5627-5637 5653 5674 5691 5693 5711 5724 5726 5733 5748 5757 5772 5778 5794 5817 5874 5894 5902 5904 5906-5907 5909 5911 5916-5919 5923-5924 5927 5929 5932 5938 5941 5948 5957 5959-5960 5962 5964

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			5979 6005 6020 6027 6037-6038 6052 6054 6057 6060-6062 6064 6068 6112 6140 6195 6225 6246 6287-6288 6306-6307 6336 6354 6371 6374 6423 6430 6451 6458 6461 6469 6478 6496 6508 6538 6545 6563 6580 6583 6595 6598 6607 6629 6638 6641-6643 6645-6675 6709 6726 6730 6765-6766 6788 6792-6793 6795 6797-6798 6800-6801 6808 6812 6836 6843 6850-6851 6855 6860 6864 6870 6875 6888 6908 6924 6950 6958 6968 6998 7003 7018 7045 7049 7056 7071-7079 7095-7096 7109 7116 7118 7126 7133 7135 7137-7139 7155-7156 7166-7167 7169 7175 7178 7275 7291 7294 7329 7422 7426 7478 7608 7627 7687 7695 7716 7815 7839 7871 7893 7914 7937 7977 8022 8095 8120 8133-8134 8148-8149 8151-8167 8169-8182 8184 8186-8192 8194-8196 8241 8246 8276 8289 8298 8336 8339 8343 8345 8349 8355-8356 8361 8363 8365 8367-8368 8370 8373 8375-8378 8385 8388 8417 8496 8518 8520 8543 8558 8561 8563 8646 8667 8710 8738-8740 8786 8803 8813 8865 8926 8946 8948 8963 8970 9010 9049 9119 9128 9142 9164 9222 9264 9289 9296 9301 9317-9318 9320 9322 9335-9341 9343-9362 9365-9368 9370 9372-9373 9375-9377 9399 9455 9457 9466 9472 9475 9480 9483 9495 9526 9533 9536 9553 9556 9558 9560 9563 9567 9570 9582 9597 9601 9630 9646 9655 9671 9683 9695 9700 9703 9715 9722 9729 9733 9752 9758 9783 9843 9848 9855 9880 9936 9942 9983 10007 10027 10103 10142 10186-10217 10274 10284 10287 10296 10313-10314 10329 10331 10349 10352 10354 10400 10405 10430 10439 10496 10500 10507 10542 10621 10709-10710 10712-10717 10750 10760 10776-10779 10782 10794 10809 10816 10837 10885 10891 10895 10913 10960 10971 10975 11057 11071 11098 11142-11143 11148 11194 11246 11401 11406 11430-11453 11455-11464 11483 11504 11508 11556 11560 11589 11596 11615 11618 11626 11631 11714 11729 11830 11833 11865 11988 12000 12006 12017 12019-12020 12041 12044 12081 12092 12178 12195 12245 12256 12277 12297 12327 12363 12383 12402-12403 12405 12426-12427 12444-12455 12457-12465 12467-12473 12475-12476 12520 12546 12576 12599 12607 12611 12614-12615 12617-12618 12622 12627 12630-12632 12636 12647-12648 12650 12685 12688 12691 12693 12696 12703 12739-12740 12743 12763 12808 12830 12834 12842 12849-12858 12860-12866 12885 12895 12906 12910 12913 12915 12917 12920 12922 12929 12933 12940-12941 12954-12955 12957 12965 12977 12990 12993 12999 13021 13039-13040 13047 13051 13054-13055 13060 13062 13072-13073 13077 13082-13083 13085 13092-13095 13098 13117 13123 13136 13173 13195-13197 13202 13249 13254 13362 13377 13391 13394 13406-13409 13411-13415 13417-13420 13456 13473-13474 13477 13480 13484 13486 13488-13489 13496-13498 13500 13502 13504 13506-13507 13511-13512 13515 13521 13546 13551 13554-13555 13560 13575 13590 13597 13613 13616-13617 13622 13630 13659-13660 13670 13678-13679 13695-13696 13701 13708 13713 13726 13748 13775 13795-13796 13815-13826 13835 13838 13859-13860 13864 13867 13869 13872 13882-13885 13888-13889 13891 13893
fetal liver-spleen	Columbia University	FLS003	525 2269 2529 2627 2636 3552 3554-3555 3597 4201 4246 4253 4422 4426 4434 4438 5536 5801 8935 9536 11136 11142 11519 11626 11927 12028 12077-12079 12086 12305 12403 12424 13021 13472
fetal liver	Invitrogen	FLV001	40 51 60 598 731 1023 1048-1049 1820 1976 2111 2268-2269 2449 2483 2529 2568 2576 2614 2809 3207 3220 3468 3595-3599 3885 4020 4346 4390 4422 4434 4543 4615 4767 4833 5092 5152-5155 5270 5505 5526 5553 5581 5715 5724 5726

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			5750 5766 5788 5794 5801 5902 6125 6147 6458 6598 6624 6645 6765 6799 6805 6862 7104 7856 8070 8156 8359 8361 8432 8740 8928 8935 8972-8974 9012 9041 9128 9264 9445 9474 9485 9511 9536 9593 9597 9703 9780 9927 9943 9945-9946 9948 10007 10137 10275 10334 10529 10557 10566-10569 10776 10828 11142 11560 11587 11618 11927 12070 12086 12105-12108 12127 12218 12236 12327 12424 12511 12541-12542 12605 12746 12916 12977 12985-12986 13020-13021 13060 13064 13135 13246-13247 13472 13479 13488-13489 13497 13533 13554 13590 13659 13680-13682 13859-13860 13882 13888-13889
fetal liver	Clontech	FLV002	360 996 1134 3226 5102 9401
fetal liver	Clontech	FLV004	998 1927 2449 2627 4025 4043 4426 4438 4834 5030 5726 6730 9474 9942 10769 11142 11587 12590 12608 13889
fetal muscle	Invitrogen	FMS001	40 150 731 894 1264 1555 1613 1821-1822 1883 1959 2161 2343 2494-2496 2555 3203 4235-4237 4434 4671 4694 4783 4885 4940 4985 5037 5045 5481 5716-5718 5724 5788 5902 5932 6624 6742-6743 6825 7054 7104 7141 7632 7800 8279-8281 8863 8935 9204 9264 9446-9449 9451 9474 9511 9536 9556 9722 9780-9781 10097 10273-10274 10330 10738-10739 10750 11098 11406 11528-11530 11560 12305 12544-12546 12590 12609 12618 12663 12670 12702 12735 12891-12892 12916 13051 13053 13195 13441-13442 13479 13613 13782
fetal muscle	Invitrogen	FMS002	3378 4348 4434 8932 11132 12033 12570 13782
fetal skin	Invitrogen	FSK001	60 142 235 319 641 683 800 1015 1050 1346 1774 1823-1825 2044 2099 2111 2161 2215 2223 2280 2367 2401 2408 2513 2516 2614 2812 2871 2926 3207 3356 3468 3490 3503 3548 3599 3720 3722 3885 4020 4032 4060 4109 4238 4240-4244 4246 4253 4277 4340 4353 4355 4388-4389 4391 4405 4434 4543 4562 4568 4654 4667 4671 4767 4940 4944 4947 5030 5037 5075 5163 5198 5296 5380 5481 5514 5581 5656 5674 5691 5719-5724 5748 5757 5801 5894 5902 5923 6057 6076 6078 6125 6147 6215 6336 6374 6409 6437-6438 6551 6563 6744 6783 6803 6805 6815 6873 6900 6908 6917 6985 6987 7037 7054 7058 7067 7075 7275 7355 7431 7484 7507 7587 7627 7642 7653 7687 7871 7977 8110 8134 8186 8226 8235 8251 8262 8282-8287 8298 8356 8365 8368 8452 8455-8456 8539 8597 8648 8679 8813 8863 8898 8921 8935 8939 8948 8951 8957 8983 9002 9041 9061 9122 9128 9257 9285 9414 9452-9453 9466 9474 9484 9503 9517 9529 9553 9563 9671 9677 9703 9714 9781 9841 9927 10007 10175 10296 10343 10491 10741-10742 10778 10795 10907 10948 11239 11513 11531 11573 11606 11626 11807 11869 11952 12000 12006 12160 12190 12218 12256 12327 12363 12371 12375 12378 12426 12438 12465 12470 12488 12549-12550 12570 12604 12615 12617 12663 12670 12735 12749 12754 12830 12893 12904 12910 12916-12917 12976-12977 13038 13084 13090 13116 13249 13254 13367 13389 13391 13443-13445 13456 13472 13479-13480 13494 13496 13499 13505 13512 13516 13551 13554-13555 13575 13590 13613 13630-13631 13644 13670 13713 13782 13784 13793 13803 13858 13866 13869 13882 13891
fetal skin	Invitrogen	FSK002	1004 3544 4834 5523 9922 9942 11134-11135 12570 13495 13499 13793 13884 13889
fetal spleen	BioChain	FSP001	997 5030 9466 11108 12033 12749 13590
umbilical cord	BioChain	FUC001	60 83 89 142 166 567 609 760 997 1302-1304 1306-1307 1309 1671 1697 1724 1848 1917 1978 2111 2154 2161 2207 2215 2315 2343 2444 2569 2576 2591 2597 2652 2866 2926 3468 3526 3599 3602 3625 3697 3722 3799-3805 3813 3839 3885 4025 4060 4173 4197 4246 4277 4340 4364 4387 4391 4395-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			4396 4405 4410 4434 4543 4589 4667 4694 4767 4834 4865 4940 4957 4963 4985 5030 5093 5096 5108 5163 5171 5213 5313 5340 5342-5344 5380 5481 5521 5526 5627 5656 5674 5714 5743 5750 5757 5801 5895 5902 5912 5938 6027 6061 6082 6125 6186 6195 6218 6241 6313 6321 6410 6458 6486-6495 6595 6607-6608 6668 6745 6825 6860 6870 6979 7054 7141 7260-7261 7275 7329 7355 7373 7383 7444 7579 7586 7677 7716 7807 7930 7935-7936 7995 8077 8093 8226 8246 8262 8298 8343 8345 8358 8387 8452 8535 8563 8635 8669 8711 8819 8858 8863 8921 8935 8943-8944 8948 8963 9001-9002 9012 9097 9130 9132 9134-9138 9264 9303-9304 9313 9332 9401 9424 9466 9472 9474 9544 9597 9665 9677 9700 9722 9734 9758 9834 9841 9848 10050 10175 10183 10267 10277 10311 10337 10446 10470 10508 10529 10566 10691 10772 10828 10961 11108 11132 11145 11239 11241 11273-11276 11293 11408 11484 11513 11596 11604 11607 11620 11695 11823 11909 11927 11942 12000 12077 12160 12245 12255-12256 12292 12305 12308 12363 12383 12404-12405 12426 12438 12511 12570 12601 12605 12630 12700 12706 12754 12875 12913 12916 12935 12957 12993 13002 13009 13020-13021 13062 13135-13136 13141 13254 13283 13318-13319 13336 13394 13400 13472 13479 13481 13488-13489 13495-13496 13499 13507 13511-13512 13516 13551 13555 13568 13575 13590 13613 13630 13644 13713 13740-13742 13866-13867 13882 13884 13888
fetal brain	GIBCO	HFB001	51 70 211-213 215-222 224-229 445 586-588 598 608 703-710 712-716 1068 1148 1178 1184 1193 1308 1516 1556 1774 1778 1957 1974 1976-1977 2010 2025-2029 2134-2135 2145 2161 2171 2174 2220 2223 2257 2266 2283 2444 2449 2510 2555 2563 2569 2590 2597 2614 2617 2717 2719-2723 2726 2729-2734 2809 2997 3002 3081 3083 3086 3109 3226-3237 3239 3257 3272 3278 3282-3283 3286 3356 3419 3460 3479 3492 3495 3544 3548 3557 3590 3604-3605 3625 3627 3634 3673 3689 3697 3709 3722 3797 3810 3839 3885 4006 4011 4019 4022 4025 4040 4054 4059 4095 4194 4201 4253 4277 4282 4355 4364 4383 4399 4405 4410 4412 4434 4473 4480 4498-4504 4507-4508 4517 4543 4654 4689 4694 4730-4733 4735 4755 4783 4820-4823 4825-4830 4845 4885 4940 4949 4957 4962 4972 4985 5017 5019 5022 5031 5037 5108 5132 5167 5171-5172 5262 5380 5481 5521 5523 5527 5553 5616 5656 5711 5740 5788 5801 5894 5902 5906 5908 5923 5938 5979-5989 6057 6082 6096 6125 6147 6159 6161-6163 6165 6207 6212-6213 6215-6222 6241 6257 6288 6336 6369 6374 6419 6422 6512 6551 6555 6595 6598 6607 6711 6767 6788 6823-6825 6860 6871 6879 6892-6894 6900 6909-6910 7058 7118-7119 7169 7220 7222 7224-7225 7228-7229 7231-7233 7275 7426 7431 7444 7481-7483 7485-7488 7560-7567 7569 7608 7743 7768 7792 7814 7845 7864 7930 7977 7995 8093 8095 8110 8140 8226 8235 8262 8345 8358 8361 8370 8372 8387 8441-8443 8445 8447 8450-8455 8520 8535 8558 8597 8603 8648 8663-8666 8677 8702 8742-8749 8811 8838-8839 8943 8951 9010 9092 9134 9137 9257 9287 9304 9317 9455 9457 9466 9472 9475 9484 9510 9515 9533 9553 9567 9644-9645 9647-9649 9734 9758 9781 9785-9786 9791 9832-9837 9927-9928 9939 9942 9970 10053 10175 10275 10277 10296 10329 10375-10378 10434 10464 10486 10496 10645 10679 10691 10778 10782 10791 10838-10844 10928 11019-11025 11027 11032 11055 11061 11108 11132 11145 11153 11208 11239 11343 11483 11513 11588 11596 11604 11606 11620-11621 11668 11691-11695 11818 11867 11869 11929-11930 12006 12033 12039 12041 12044 12047 12066 12078 12197 12218

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			12327 12375 12405 12411 12424 12438 12521-12522 12564 12570 12576 12582 12590 12599 12601 12604 12611 12626 12630 12637-12640 12670 12674 12688 12694-12695 12703 12734 12748 12754 12785 12835 12840 12913 12916-12917 12929-12930 12935 12940 12946 12959 12989 12993 13020 13051-13052 13054 13073 13083 13090 13108-13110 13117 13131 13136 13148 13158 13160 13169 13184-13186 13254 13263 13277 13295 13389 13394 13438 13456 13477 13479-13481 13489-13490 13494-13500 13502-13503 13506 13512 13516 13535-13536 13555 13575 13590 13592 13605-13606 13613 13616 13627 13630 13644-13645 13660 13670 13679 13687 13713 13793 13803 13838 13861 13866 13868-13869 13875 13888 13891 13896
macrophage	Invitrogen	HMP001	1002 2253 3548 4011 4058 4201 4246 4282 5526 8093 8262 11137 12039 12426 12511 12521 13888
infant brain	Columbia University	IB2002	6 89 142 211 276 307 518 589-590 598 644 717-718 720-721 773 841 921-924 926-932 937 1023 1091 1147-1150 1289 1493 1522-1525 1557-1564 1724 1778 1928 1947 1976 1978 2097 2111 2138 2172 2218-2223 2257 2283 2401 2418-2419 2444 2449 2516 2555 2563 2569 2575 2597 2627 2809 2818 2880 2932 3087 3166 3226 3241-3245 3272 3435 3474 3476-3477 3481 3483-3486 3488-3489 3492 3494-3495 3497 3538 3544 3548 3572 3595 3597 3669 3671-3672 3709 3720 3722 3737 3797 3817 3885 3985 3997-3999 4006 4011 4022-4031 4040 4095 4201 4246 4253 4277 4297 4355 4405 4410 4434 4676 4689 4767 4775 4831-4834 4837 4885 4962 4983 4998 5017-5019 5021-5031 5035-5037 5040 5095 5143 5164 5167 5169 5233-5236 5380 5505 5517-5519 5523 5536 5546-5548 5691 5695 5724 5750 5766 5788 5801 5895 6027 6048 6082 6095 6111 6119 6132 6135 6147 6223 6301-6303 6305-6307 6374 6411 6459 6555 6563 6591-6595 6601 6606-6607 6609-6611 6613-6614 6624 6711 6767 6799 6805 6909 6946 7018 7052 7067 7118 7133 7166 7168-7169 7386 7464 7469 7473 7571-7573 7697-7698 7716 7814 7827-7828 7905 7934 7977 8071-8073 8084 8098-8104 8235 8320 8331 8351 8354-8356 8364-8365 8370 8372 8376 8452 8520 8572 8667-8668 8750-8752 8813 8830 8863 8883-8885 8950 8958 8963 9040-9041 9251 9257 9267 9293-9298 9313 9424 9454 9456-9457 9466 9510 9514-9515 9533 9542 9556 9576-9577 9597 9626 9646 9722 9758 9787 9903-9904 9988 10027 10138-10139 10155-10160 10284 10296 10319 10346 10496 10536-10542 10557 10670-10672 10680 10682 10778 10791 10897 10971 11028-11030 11052 11100-11109 11122 11132 11192 11392 11403-11404 11447 11513 11568 11595 11606 11626 11638 11818 11927 11931-11933 12028 12030-12031 12033-12034 12036-12039 12041 12047 12049 12146 12155-12157 12160 12327 12344 12363 12383-12385 12402 12405-12408 12424 12439 12521-12522 12570 12590 12601 12611 12615-12616 12626 12639 12684-12685 12688 12696-12697 12707 12725-12727 12729 12748-12749 12754 12763-12764 12830 12904 12913 12916-12917 12923 12929 12959 12977 12990 12996 13000 13047 13051 13054 13062 13073 13082-13084 13117 13141 13187-13188 13196 13218-13219 13249 13257 13277 13284-13285 13351 13389 13394-13395 13456 13458 13479-13481 13488 13494-13495 13499 13503 13512 13516 13530 13535 13575 13607-13608 13613 13616 13641 13649 13659 13679 13705-13706 13708 13713 13793 13798-13801 13803 13861 13869 13872 13875 13883-13884 13888 13891-13893 13896
infant brain	Columbia University	IB2003	6 46 746 1914 1947 1959 1994 2111 2220 2257 2516 2962 3166 3226 3272 3435 3572 3885 4277 4377 4410 4833 4837 5029-5030 5040 5147 5259 5851 6147 6423 6595 6598 6611

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			7145 7169 7716 8262 8354-8355 8364 8452 8863 9264 9457 9515 9556 9577 9626 9646 9715 9781 9928 9970 10027 10329 10376 10557 10791 11083 11289 11406 11459 11513 11663 11933 12039 12405 12546 12570 12601 12616 12685 12754 12913 12916 12954 12977 12990 13000 13056 13062 13195 13257 13392 13456 13458 13480-13481 13488 13497 13499 13506-13507 13511-13512 13514 13516 13549 13616-13617 13649 13793 13861 13884 13888 13891 13896
infant brain	Columbia University	IBM002	1564-1565 1976 2218 2420 3166 4391 5788 6147 6302 10329 11108 11513 12039 12729 12830 13062 13257 13512 13617 13803 13859 13861
infant brain	Columbia University	IBS001	927 931 1564 2915 3166 3737 3885 5029 5044 5095 5143 5701 5894-5895 6135 6307 7018 8452 8963 9424 9556 9626 10296 10542 11083 11100 11108 11167 12383 12611 12785 13257 13280 13479 13495 13506 13617 13679 13713 13896
lung, fibroblast	Strategene	LFB001	67 235 552 641 746 1346 2090 2186 2343 2770 2812 2871 3203 3537 3548 3634 3812 3839 4073 4253 4434 4654 4783 4786 4821 4940 4957 4963 4985 5163 5172 5492 5523 5788 5911 5938 6027 6048 6057 6223 6272 6322 6599 6873 6886 6997 7204 7759 7815 7933 7977 8226 8235 8262 8354 8452 8648 8742 8791 8830 8863 8898 8944 8948 8951 9000 9076 9097 9128 9290 9304 9414 9455 9466 9472 9476 9503 9544 9563 9671 10000 10050 10342 10376 10434 10447 10679 10760 10857 11055 11094 11289 11402 11607 11818 11837 12000 12006 12178 12363 12405 12462 12582 12617 12636 12670 12754 12913 12916-12917 12925 12940 13126 13377 13394 13446 13456 13494 13497 13499 13502 13575 13613 13670 13679 13691 13795 13797 13867-13868 13882
lung tumor	Invitrogen	LGT002	51 69 88 142 158 211 444 598 608 722-723 733 795-797 799 1004 1023 1135-1145 1308 1338 1346 1493 1526-1533 1535 1566-1569 1626 1654 1816 1841 1848 1860 1976 1983 2090 2150 2161 2173 2223 2257 2266 2296-2297 2342 2402 2449 2483 2555 2569 2576 2591 2611 2623 2724 2809 2868 2880 2926 3088 3166 3272 3347 3571 3658-3666 3673 3722 3737 3797 3839 3885 4000-4002 4006 4011 4025 4032-4034 4060 4133 4201 4246 4253 4277 4282 4340-4341 4384 4388 4391 4399 4402 4405 4434 4543 4622 4671 4686 4767 4783 4791 4833-4834 4836 4885 4903 4940 4957 4987 4998 5001 5017-5018 5030 5037 5052 5108 5163 5210 5227-5232 5422 5451 5520-5528 5581 5627 5637 5674 5693 5713 5724 5748 5766 5772 5788 5894-5895 5902 5907 5917 5938 5990 6005 6020 6094 6129 6147 6171 6224-6225 6239 6241 6253 6322 6336 6404-6410 6461 6508 6595-6598 6624 6765 6793 6798 6851 6870 6873 6994 7003 7008 7045-7047 7053-7055 7067 7085 7109 7116 7139 7141 7158 7169 7187 7468 7517 7524 7579 7608 7676 7687 7716 7761 7825-7826 8074-8078 8105-8106 8139 8235 8298 8323 8335 8345 8354 8359 8363 8365 8370 8372 8375 8456 8563 8633 8648 8678 8741 8792 8798 8805 8831 8863 8870 8989 9002 9010 9012 9032-9037 9039 9051 9076 9080 9092 9128 9257 9269-9271 9273-9274 9276 9299-9301 9399 9414 9424 9457 9466 9472 9480 9484-9485 9517 9533 9536 9539 9544 9567 9612 9626 9665 9677 9689 9700 9703 9758 9773 9784 9788 9838-9841 9870 9928-9929 9942 9970 9983-9986 10140-10143 10149 10161 10175 10275 10296 10319 10330 10346 10349 10386 10496 10508 10561 10591-10593 10673-10674 10742 10774 10778 10782 10788 10802 10836 10971 11031-11033 11064 11108 11132 11182-11191 11393-11395 11406 11459 11462 11560 11565 11588 11596 11604 11606-11607 11615 11618 11629 11821 11823 11845 11869 11932 11934 11961 11978-11980 12000 12006 12017 12022 12033 12078 12080 12127 12150-12154 12178 12236

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			12327 12363 12371 12379 12386-12393 12409 12424 12521 12570 12590 12616-12618 12630 12637 12674 12691 12693 12698-12699 12716 12749 12754 12760 12762 12830 12832-12833 12842 12904 12907 12913 12917 12923 12993 13020 13031-13032 13035 13060 13062 13064 13072 13083 13090 13136 13150 13193 13197 13202 13234 13277 13279 13281-13282 13295-13296 13381-13383 13385-13386 13394 13400 13456 13477 13479-13480 13484 13488-13489 13492 13494 13496-13497 13499 13503 13506-13507 13511 13554 13575 13597 13616 13630 13646 13649 13659-13660 13670 13701-13704 13713 13743 13782 13788-13789 13791-13792 13802-13803 13864 13867-13868 13872 13882-13884 13888 13892
lymphocytes	ATCC	LPC001	142 316 1034 1977 2161 2342 2367 2483 2513 2529 2932 2962 3272 3711 3722 3839 4006 4277 4282 4405 4434 4836 4970 5037 5239 5451 5706 5796 5924 5941 6048 6147 6301 6765 6793 6805 6815 7133 7171 7275 7320 7687 7815 8033 8093 8095 8105 8110 8246 8262 8345 8365 8368 8370 8535 8664 8674 8722 8834 8858 8898 8919 8963 9012 9257 9456-9457 9484 9539 9544 9560 9758 9928 9939 10175 10201 10284 10333 10496 10520 10529 10679 10744 10766 11032 11108 11508 11513 11519 11853 11862 12033 12160 12363 12387 12402 12424 12438 12520 12560 12570 12663 12830 12913 12923 12968 13021 13066 13072 13085 13140 13147-13148 13394 13479 13488 13495 13499 13507 13575 13867 13869 13888
leukocyte	GIBCO	LUC001	21 49 51 67 83 88 94 142 211 316 326 340 368 371 403 438 443-455 464 485 518-523 525-530 532-535 591-592 616 724-726 744 800 841 889 943 1004 1346 1556 1570-1572 1826 1914 1916 1927 1959 1976-1977 2095 2103-2107 2124-2127 2134 2138 2145 2161 2163 2175 2215 2223 2237 2266 2343 2363 2412 2421 2449 2483 2529 2555 2569 2575 2597 2617-2618 2695-2696 2827 2876-2877 2899 2903 2961 2965-2975 3002 3035-3038 3040-3046 3109 3132 3166 3199 3203 3207 3246-3251 3265 3272 3286-3287 3312 3344 3356 3361 3378 3400 3434 3451 3468 3479 3484 3494 3503 3552 3590 3610 3627 3634 3704 3709 3720 3722 3725 3757 3797 3804 3810 3839 3885 3985 3996 4006 4009-4011 4019 4022-4023 4025 4035-4037 4058 4060 4194 4253 4269 4277 4297 4341 4362 4388 4390-4391 4396 4399-4400 4402 4405 4410 4412 4431 4434 4534 4543 4594 4615 4630 4641 4651-4663 4670 4688-4689 4694 4699-4706 4736 4755 4758-4759 4767 4775 4783 4791 4798 4828 4835-4840 4845 4865 4930 4940 4947-4948 4957 4962 4964 4972 4976 4985 4991 4998 5001 5003 5017 5022 5030 5037 5040 5075-5076 5108 5118 5143 5163 5171-5172 5313 5481 5503 5505 5521 5523 5526 5535 5549 5552 5656 5691 5724 5726 5740 5750 5766 5772 5788 5794 5796 5801 5865 5874 5894-5895 5906 5908 5923-5924 5928 5938 5989 6027 6042 6057 6063-6064 6082 6094-6103 6125 6130-6137 6142 6147 6166-6167 6171 6181 6214 6226 6239 6253 6301 6307 6371 6374 6418 6512 6662 6716 6730 6788 6792 6799 6815 6836 6860 6873 6876-6877 6886-6887 6945 6975 6979 7018 7037 7056 7058 7067 7116 7118 7137 7155 7158 7171 7275 7315 7384-7385 7387-7389 7391-7393 7395-7398 7452-7463 7468 7494 7608 7671 7676 7687 7714 7733 7792 7815 7845 7864 7870 7905 7930 8093 8107 8110 8120 8139-8140 8224 8226 8262 8276 8320 8363-8365 8368 8375 8387 8432 8452 8456 8520 8535-8536 8539 8562-8563 8577 8579 8597 8599-8601 8603-8605 8608 8640-8646 8648 8664 8669 8674 8677 8691 8702 8722 8753 8755 8798 8815-8816 8830 8838-8839 8858 8863 8870 8876 8898 8943-8944 8948 8951 9001 9010 9012 9061 9076 9092 9097 9128 9257 9304 9454-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			9456 9466 9472-9475 9484 9497 9511 9515 9517 9529 9536 9542 9556 9567 9612 9626 9702-9703 9723 9731-9738 9745 9752 9758 9762-9770 9780-9781 9789 9791 9808 9832 9842 9927-9928 9939 9942 9970 10000 10175 10275 10319 10329-10331 10333 10345-10346 10412 10428 10430 10433 10435-10436 10453-10454 10456 10465 10470 10475 10488 10496 10513 10566 10645 10679 10683 10691 10774 10776 10778 10788 10791 10795 10802 10891 10928 10930 10935-10941 10948 10961-10962 10964-10967 11034 11055 11108 11132 11136 11145 11239 11289 11405-11406 11459 11483 11519 11588 11604 11606 11618 11620 11693 11695 11805 11809-11816 11818 11830 11833 11848 11850-11856 11868 11927 11932 11934-11935 11937-11939 11954 11988 12000 12006 12009 12028 12033 12037 12039 12047 12053 12092 12097 12118 12127 12143 12159-12160 12178 12186 12327 12336 12363 12375 12387 12402 12404-12405 12410 12424 12438 12442 12490 12511 12521-12522 12556 12560 12564 12570 12590 12599 12601 12605 12615-12617 12636 12643 12653 12660 12670 12672-12680 12693 12700-12701 12706-12707 12715 12735 12754 12823 12895 12904 12910 12913 12916-12918 12923 12925 12935 12945 12947 12955-12956 12972 12977 12993 13002 13020-13021 13051 13062 13066 13072-13073 13082 13085 13117 13126 13130-13131 13135 13138-13140 13147-13151 13161 13169 13189 13254 13295 13391 13394 13425 13436 13442 13456 13477 13479 13488-13490 13494-13496 13498-13503 13505-13507 13512 13530 13546 13550-13551 13554 13557-13564 13575-13581 13590 13592-13593 13609 13613-13616 13627 13630 13645-13646 13660 13679 13713 13782 13803 13864 13866-13869 13872 13883 13885 13888 13893
leukocyte	Clontech	LUC003	536 539 541-543 593 728 1552 1927 1974 3089 3252-3253 3434 3548 3709 3711 3722 3797 4011 4019 4040 4060 4194 4201 4277 4282 4622 4707 4791 4841-4842 4949 4998 5001 5030 5345 5895 6052 6138 6227 6419 6595 6754 6765 6767 6788 6911-6912 7049 7139 7171 7464 7575-7577 8093 8110 8116 8365 8370 8375 8592 8648 8830 8863 8944 9466 9544 9612 9758 9771 9773 9790 9929 10326 10346 10679 10779 10961 11035-11036 11132 11401 11513 11853 11940 12160 12256 12405 12670 12674 12681 12840 12904 12906 13020 13051 13162 13280 13400 13488 13511 13516 13554 13670 13679 13860 13869 13883 13889
melanoma from cell line ATCC #CRL 1424	Clontech	MEL004	83 142 360 447 841 1061 1346 1516 1573 1724 1959 1963 2159 2175 2343 2367 2513 2652 2812 2876 2961 3001 3272 3345 3604 3627 3722 3754 3839-3840 3985 3988 4040 4194 4246 4341 4515 4568 4641 4667 4862 4940 5075 5132 5163 5481 5492 5523 5740 5788 6008 6112 6119 6147 6189 6195 6272 6287 6418 6423 6508 6563 6726 6860 6870 7134 7166 7169 7275 7294 7579 7677 7748 7831 7852 7980 8110 8116 8226 8235 8262 8320 8345 8363 8370 8456 8563-8564 8633 8838 9001-9002 9128 9296 9304 9313 9484 9542 9544 9646 9703 9758 9780-9781 9808 9999 10007 10027 10296 10346 10470 11145 11239 11289 11401 11406 11508 11596 11606 11620 11823 11899 11950 12019 12107 12160 12292 12329 12363 12405 12436 12523 12599 12658 12670 12749 12754 12774 12842 12930 13020 13296 13394 13458 13489 13498 13501 13507 13551 13554 13575 13613 13616 13649 13660 13743 13775 13868
mammary gland	Invitrogen	MMG001	51 67 142 449 594 598 616 708 729-733 1004 1060 1194-1196 1310-1315 1329 1536-1541 1826 1848 1858 1914 1947 1974 1976-1978 2174-2175 2283 2299 2310 2336 2343 2403-2406 2449 2483 2555 2563 2576 2597 2611 2614 2617 2812 2827

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			2926 3191 3207 3490 3503 3673 3714-3715 3722 3737 3777 3796 3807-3819 3885 4003-4004 4060 4330 4346 4355 4391 4399 4434 4543 4671 4767 4791 4836-4837 4843 4940 4964 4975 4985 4998 5001 5017 5019 5022 5052 5108 5155 5163 5168-5169 5269 5345-5347 5466 5526-5527 5529 5581 5615 5724 5772 5778 5788 5794 5894-5895 5902 5907 5911 5917 5923-5924 6048 6057 6064 6094 6109 6125 6147 6156-6157 6215 6228 6321 6430 6458 6496-6499 6545 6579 6598-6599 6607 6711 6765 6780 6793 6797 6799 6805 6880 6931 6979 6987 6996 7012-7013 7048 7054 7057 7104 7141 7158 7169 7517 7578-7581 7583 7872 7937-7940 8110 8139 8156 8224 8262 8292 8298 8336 8359 8363 8365 8372 8452 8619 8628 8646 8648 8722 8758-8759 8782 8813 8815 8863 8882 8939 8963 8983 9012 9097 9111 9139-9141 9143 9157 9164 9264 9313 9327 9347 9466 9484 9511 9517 9544 9553 9563 9577 9646 9700 9703 9734 9746 9780 9791 9844-9846 9927-9928 9942 10006-10007 10051-10055 10144 10146-10147 10162 10285 10330 10346 10439 10490-10491 10496 10542 10557 10600 10628 10645 10691 10729 10772 10796 10983 11038 11108 11132 11216-11217 11277-11279 11396-11397 11406 11459 11513 11596 11606-11607 11626 11628 11823 11830 11833 11902 11941-11943 11979 12000 12009 12037 12047 12078 12093 12146 12160 12181 12258-12263 12281 12327 12371 12383 12387 12395 12424 12426 12438 12442 12484 12511 12564 12570 12590 12594 12601 12605 12609 12615-12616 12618 12621 12653 12670 12691 12699 12702-12703 12749 12771 12791-12794 12834-12835 12885 12898 12905 12913 12916-12917 12920 12923 12955 12976-12977 13002 13020 13035 13047 13062 13064 13066 13072-13073 13084 13090 13092 13127 13135 13169 13190 13193 13195 13249 13254 13296 13320-13322 13392 13456 13477 13479-13480 13488-13489 13494 13497-13499 13505-13507 13512 13514 13516 13546 13551 13554-13555 13590 13611-13612 13631 13649 13670 13713 13743-13745 13838 13864 13866-13867 13869 13872 13875 13882-13885 13888 13891
induced neuron cells	Strategene	NTD001	88 1493 1552 1561 2034 2090 2510 2522 2570 2576 2623 3226 3272 3839 4006 4058 4282 4384 4694 4783 4888 5103-5104 5172 5523 5541-5542 5684 5766 5788 5794 5938 6082 6195 6418 6449 6607 6960 6972 7058 7494 7815 7937 8370 8929-8930 8996 9000 9003 9128 9157 9289-9291 9472 9928 10275 10550 10729 10760 11607 11736 11818 12039 12075 12094 12256 12363 12381 12423 12522 12707 12735 12863 12906 12958-12959 12963 13020 13052 13054 13196 13202 13310 13368 13378 13438 13479 13486 13506-13507 13546 13616 13654 13659 13691 13795 13891
retinoid acid induced neuronal cells	Strategene	NTR001	1552 3839 3885 4282 4434 5030 9308 9466 11108 11132 12034 12039 12405 12424 12570 12905 13047
neuronal cells	Strategene	NTU001	88 708 1552 1561 1947 2223 2490 2614 3272 3610 3625 3627 4032 4201 4253 4297 4337 4434 4543 4622 4641 5104 5163 5523 5541 5684 5788 6094 6147 6307 6601 6862 7755 8648 8740 8996 9097 9128 9289 9308 9466 9511 9544 10333 10779 10963 11032 11108 11132 12034 12039 12075 12405 12424 12570 12590 12706 12749 12830 12905 12917 12963 13035 13073 13280 13490 13507 13613 13616
pituitary gland	Clontech	PIT004	1148 5674 6068 6371 7759 8093 10430 12605 12636 12658 12916 13021 13082 13875
placenta	Clontech	PLA003	3885 4025 5030 5536 5711 5923 8349 8367 9921 12033 12076 12403 12511 13655 13885
prostate	Clontech	PRT001	67 83 142 225 235 379 486 572 616 1286 1316-1329 1538

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			2023 2139 2147 2161 2266 2290 2337-2344 2346 2367 2422 2444 2511 2631 2809 3001-3002 3084 3419 3563 3625 3627 3797 3810 3821-3827 3829-3835 4019 4022 4173 4282 4332 4434 4667 4755 4791 4862 4865 4998 5001 5066 5171 5349 5352-5356 5358-5361 5363-5364 5481 5581 5656 5674 5723 5740 5902 6027 6047 6125 6321 6371 6374 6427 6458 6463 6500-6507 6509 6563 6598 6643 6793 6803 6871 7014-7015 7058 7104 7275 7320 7608 7635 7733 7842 7852 7864 7941-7943 7946-7951 7953-7954 8077 8087 8093 8110 8224 8226 8452 8487 8520 8558 8635 8677 8863 8950 8963 8983 8997 9012 9145-9148 9150-9154 9269 9302 9317 9466 9503 9528 9646 9703 9780 10000 10027 10056-10060 10491 10629 10691 10777 10891 11145 11200 11239 11281 11283-11295 11344 11406 11761 11837 11862 12006 12166 12264-12278 12305 12363 12368 12411-12412 12438 12467 12685 12691 12729 12734 12795-12798 12800 12830 12863 12904 13010-13011 13104 13146 13295 13323-13326 13377 13394 13456 13473 13477 13489 13516 13533 13550 13611 13678-13679 13746 13866-13867 13884 13889
rectum	Invitrogen	REC001	6 67 142 683 731 997 1178 1909 1959 2005 2023 2596 2611 2614 2809 2926 3314 3333 3455 3722 3780 3870 4173 4355 4837 4949 4987 5526 5580 5615 5674 5691 5724 5788 5895 5909 5924 6057 6112 6195 6765 6805 7815 7833 8095 8664 8863 9517 9539 9544 9884 9927 10400 10666 11401 11513 11606 11985 12160 12327 12428 12693 12848 12910 12977 13051 13064 13072 13420 13494 13497 13507 13512 13515-13516 13866 13869 13884 13888
salivary gland	Clontech	SAL001	67 731 800 997 1054 1914 2267 2395 2529 3136 3595 3627 4011 4192 4246 4330 4434 4641 4957 4987 5040 5052 5163 5451 5481 5706 5723 5788 5895 6219 6621 6801 6900 6975 7045 7733 8110 8372 8535 8563 8635 8830 8951 9000 9010 9051 9313 9472 9475 9671 9724 9758 9927 10027 11145 11695 11725 12017 12284 12363 12424 12427 12570 12609 12670 12674 12693 12977 13035 13307 13554 13617 13867 13872 13889 13891
salivary gland	Clontech	SALs03	1516 1724 1858 5030 6186 13657 13864
skin fibroblast	ATCC	SFB001	2251-2252 5788 6068 12511
skin fibroblast	ATCC	SFB002	6068 8951 12511
skin fibroblast	ATCC	SFB003	4025 5895 7741
small intestine	Clontech	SIN001	142 319 627 654 1034 1063 1197-1198 1330-1338 1340-1359 1575 1646 1774 1814 1978 2161 2347-2354 2409 2876 3046 3419 3460 3605 3716-3718 3737 3797 3837-3839 3841-3843 3845-3857 3885 3986 4060 4201 4301 4351 4385 4568 4689 4694 5076 5163 5270-5273 5304 5326 5365 5367-5372 5374 5503 5550 5701 5772 6064 6094 6171 6288 6427 6430-6432 6438 6510-6522 6598 6615 6793 6815 6997-6998 7016-7018 7054 7058 7072 7309 7450 7604 7769 7811 7873-7876 7955 7957 7959-7962 7964 8120 8298 8350 8452 8830 8863 8950-8951 8966 9010 9073-9075 9119 9126 9128 9155-9166 9303 9544 9560 9780 9884 9928-9929 10008-10010 10061-10068 10097 10262 10330 10351 10601 10630-10634 10760 10983 11061 11219 11296-11308 11310-11313 11513 11620 11693 12182-12183 12280-12287 12327 12363 12488 12707 12799-12801 12922 12991 13012-13014 13035 13051 13064 13297 13307 13328-13332 13335 13382 13499 13506 13554 13560 13575 13631 13695 13714 13747-13749 13751 13882 13884
skeletal muscle	Clontech	SKM001	1104 1346 2363 2367 2495 2555 2876 2880 3555 3634 3722 4011 4022 4194 4201 4253 4277 4282 4434 4641 4940 4972 4998 5343 5481 5523 5801 6005 6336 6873 7408 7995 8110 8120 8235 8262 8292 8345 8372 8576 8740 8830 8936 8951

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			9303 9689 10616 10679 11132 11145 11825 11869 11927 12387 12442 12467 12570 12663 12735 12916 13280 13503 13872 13885
skeletal muscle	Clontech	SKM002	8535
skeletal muscle	Clontech	SKMs03	6336 6962 8936
skeletal muscle	Clontech	SKMs04	770 1724 3797 4277 6336 12405 13658
spinal cord	Clontech	SPC001	83 142 390 415 598 668 708 731 1184 1199-1207 1360-1375 1377-1396 1516 1574-1576 1595-1596 1849 1927 2070 2129 2161 2311-2314 2345 2355-2368 2423-2424 2430 2484 2529 2569 2576 2876 3215 3249 3272 3283 3532 3584 3627 3634 3711 3719-3722 3737 3839 3860-3884 4011 4025 4038-4039 4043 4055-4056 4173 4246 4282 4354 4375 4391 4434 4681 4767 4781 4808 4964 4985 4998 5037 5163 5233 5274-5277 5375-5392 5394-5402 5523 5569-5570 5581 5615 5723 5788 5835 5902 5928 5936 6047 6078 6082 6211 6288 6374 6433-6435 6512 6523-6531 6534 6595 6616 6625-6626 6788 6894 6979 6999 7018-7026 7126 7166 7359 7473 7642 7653 7807 7814 7877-7879 7965-7968 7970 7972-7980 8105 8108-8110 8139 8246 8298 8345 8363 8368 8482 8603 8646 8884 8898 8981 9010 9012 9076-9078 9098 9167-9179 9184-9189 9264 9302 9304 9319-9320 9455 9466 9520 9530 9544 9556 9567 9781 9895 9901 9928 9942 9947 9969 9999 10007 10069-10077 10079-10085 10177 10296 10326 10346 10376 10422 10566 10602 10635-10638 10679 10685-10686 10729 10776 11132 11220 11246 11314-11323 11325-11330 11417-11418 11459 11513 11818 12000 12011 12017 12033 12039 12160 12184-12185 12288-12292 12295-12299 12301-12305 12363 12375 12383 12387 12402 12413 12442 12468 12527 12605 12617 12636 12657-12658 12739-12740 12754 12772 12802-12809 12830 12835 12841-12842 12905 12923 12940 12976 13003 13015 13017-13021 13051-13052 13117 13126 13136 13260 13277 13283 13295 13336-13343 13367 13442 13456 13473 13477 13481 13495 13497 13499-13500 13507 13516 13659 13670 13713 13715-13716 13748 13752-13759 13803 13869 13872 13884-13885 13888 13893 13896
adult spleen	Clontech	SPLc01	800 1927 4032 4834 6064 6135 6195 6446 6788 6873 7166 7455 8966 9929 10744 12402 12564 12590 12691 12904 12933 13082 13500 13506-13507 13516 13575 13864 13869 13883 13889
stomach	Clontech	STO001	21 83 142 1004 1208-1215 1217-1219 1397 1399-1405 1671 2315-2316 2345 2369-2373 2375 2575-2576 2809 2846 2984 3136 3166 3537 3610 3698 3723-3725 3839 3885-3897 4057-4059 4173 4277 4410 4480 4667 4791 4808 4940 4987 5262 5278-5281 5283-5284 5403-5405 5407-5424 5481 5656 5674 5796 5904 6418 6436-6440 6535-6540 6563 6627-6629 6765 6940 7000-7001 7027-7030 7064 7135 7509 7604 7880-7885 7981-7990 8087 8110 8120 8143 8226 8452 8535 9010 9079-9081 9191 9193-9196 9304-9306 9313 9317 9321 9715 10007 10011-10013 10086-10093 10178-10179 10603-10605 10640-10642 11069 11167 11221-11222 11331-11337 11339-11343 11419 11513 11818 12186-12190 12307-12314 12327 12363 12425-12427 12438 12617 12773-12774 12810-12811 12834 13082 13103 13298-13299 13344-13349 13592 13630 13670 13717 13760-13764 13782 13888
thalamus	Clontech	THA002	579 598 616 1065 1148 1220-1221 1223-1226 1407-1432 1597 2266 2317-2319 2340 2342 2376-2378 2380 2431 2444 2555 3093 3230 3286 3537 3722 3726-3732 3737 3898-3902 3904-

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			3918 3920-3922 4060-4062 4173 4201 4282 4360 4391 5270 5285-5288 5425-5438 5440-5449 5701 5902 5938 6137 6437 6458 6461 6541-6544 6546-6550 6630 6755 7031 7095 7119 7166 7484 7579 7815 7849 7905 7977 7991-8000 8002-8005 8126-8128 8134 8363 8558 8664 8786 8890 8930 8963 9082-9085 9128 9197-9199 9201-9216 9251 9264 9308 9503 9515 9556 9646 9703 9928 10014-10017 10027 10094-10100 10102 10267 10496 10606-10608 10643-10650 10696 10891 11145 11223-11225 11344-11355 11406-11407 11420 11513 11604 11695 11791 11950 12022 12107 12191-12199 12315-12331 12363 12375 12405 12428-12430 12688 12706 12735 12748 12754 12812-12813 12815 12835 12914 12959 13020 13023-13025 13060 13073 13300 13351-13358 13402-13403 13489 13496 13506 13512 13719-13720 13765-13768 13770 13872 13883
thymus	Clontech	THM001	51 142 150 332 346 360 438 546 731 760 895 1004 1104 1227-1234 1264 1391 1516 1577-1585 1812 1860 1877 2129 2174 2215 2263 2321-2322 2408 2414 2425 2449 2490 2555 2569 2575 2611 2847 2880 3435 3530 3722 3727 3733 3735-3742 3839 3870 4006 4041 4043-4046 4060 4109 4375 4396 4399 4434 4667 4671 4759 4791 4885 4976 4987 5108 5289-5290 5466 5481 5526 5553-5554 5796 5956 5979 6020 6186 6253 6336 6371 6411 6438 6441-6447 6617-6621 6765 6788 6797 6870 6886 6908 6972 6994 7059-7060 7126 7141 7166 7168 7310 7383 7450 7494 7632 7716 7779 7887 7889-7891 8087 8111-8114 8117-8120 8139 8226 8343 8368 8370 8452 8456 8633 8830 8898 8921 8963 9010 9076 9086-9087 9128 9134 9264 9304 9307-9313 9401 9466 9497 9544 9563 9582 9612 9626 9646 9758 9781 9927 9937 9969 10018 10065 10163-10166 10175 10284 10329 10414 10557 10609-10610 10617 10679 10687-10691 10742 10760 10772 10776 10778 10881 10891 11187 11194 11226-11229 11274 11406 11408-11410 11412 11459 11513 11676 11695 11830 11865 11942 12000 12006 12016 12022 12092 12160 12186 12200-12204 12327 12363 12414-12417 12427 12462 12470 12490 12564 12693 12706 12735 12740 12835 12840 12898 12905 12910 12913-12914 12916 12977 13020 13036 13051 13062 13111 13126 13141 13295 13301-13305 13326 13351 13391 13396-13397 13456 13484 13498 13505 13507 13512 13516 13546 13713 13803 13805 13808 13866 13869 13885 13888
thymus	Clontech	THMc02	16 27 51 67 142 390 598 1233 1493 1508 1586-1588 1598-1603 1724 1841 1918 1927 1959 1976 1979 1989 2057 2107 2161 2223 2290 2407 2426-2428 2432-2433 2444 2487 2569 2617 3537 3664 3711 3720 3722 3737 3780 3817 4021 4025 4047-4049 4060 4063-4068 4246 4277 4377 4405 4688 4694 4759 4791 4837 4949 4957 4992 4998 5037 5052 5108 5505 5553 5555-5564 5571-5572 5683 5766 5772 5788 5796 5894 5911 6048 6186 6225 6288 6306 6333 6621 6728 6730 6765 6767 6780 6788 6815 6867 6873 6979 7049 7061-7062 7166 7169 7676 7687 7809 8084 8121 8129-8130 8179 8262 8354 8363 8365 8375 8482 8597 8654 8740 8786 8791 8963 9076 9157 9257 9264 9314 9323-9324 9401 9454 9466 9519 9529 9536 9637 9700 9703 9927-9929 9940 10004 10007 10070 10167 10169-10172 10180 10182 10267 10284 10326 10331 10383 10439 10452 10542 10605 10691 10693-10694 10697 10744 10775 10778 10795 11046 11098 11274 11413-11415 11462 11519 11576 11592 11606 11618 11621 11627-11628 11693 11807 11814 12022 12034 12044 12080-12081 12086 12160 12236 12256 12327 12335 12363 12368 12387 12418-12419 12424 12433 12556 12560 12564 12570 12577 12594 12599 12612 12663 12735 12754 12836-12839 12844 12905 12913

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			12955 12977 13020 13035 13037 13051 13062 13072 13087 13193 13280 13295 13398 13404 13456 13488 13499-13500 13503 13507 13516 13533 13535 13546 13611 13613 13630 13646 13649 13659 13670 13678-13679 13713 13769 13775 13803 13806-13808 13810-13811 13869 13872 13882-13883 13885 13888-13889
thyroid gland	Clontech	THR001	49 115 142 360 641 698 800 1004 1134 1193 1233 1235 1237-1241 1329 1433-1471 1473-1476 1478-1479 1481 1589-1591 1593-1594 1604 1606-1612 1614-1615 1639 1671 1906 1927 1959 1976 1994 2090 2099 2134 2150 2161 2174 2186 2215 2343 2367 2381-2387 2429 2434-2437 2449 2490 2510 2529 2555 2569 2662 2827 2901 2961 2997 3001 3109 3265 3286 3425 3490 3503 3560 3643 3698 3722 3737 3743-3747 3797 3839 3885 3923-3941 3943-3954 3956-3961 4009 4050-4053 4060 4069-4082 4109 4173 4194 4253 4277 4282 4348 4354-4355 4376 4391 4396 4412 4434 4641 4689 4694 4755 4759 4783 4791 4834 4957 4972 4985 4987 5018 5037 5052 5075-5076 5108 5147 5163 5171 5270 5292-5294 5317 5380 5450-5461 5464-5481 5503 5523 5565-5568 5573-5574 5693 5711 5724 5757 5772 5788 5894-5895 5917 5923 5938 5959 5966 6005 6013 6027 6047-6048 6061 6064 6125 6135 6139 6189 6215 6240 6336 6371 6374 6448-6449 6458 6508 6538 6551-6554 6556-6561 6563-6572 6595 6598 6607 6622-6623 6631-6636 6793 6803 6815 6873 6899 6955 6958 7032-7038 7040 7043 7063 7066-7067 7079 7116 7155 7233 7275 7455 7669 7743 7792 7839 7845 7857 7892-7893 7910 7930 7989 8006-8008 8010-8033 8053 8067 8087 8092-8093 8110 8116 8120 8122-8124 8131-8143 8196 8226 8235 8262 8345 8365 8368 8370 8372 8375 8452 8520 8535 8543 8597 8646 8711 8760 8816 8839 8863 8898 8930 8948 8951 8963 8966 9010 9055 9088-9091 9142 9157 9217-9220 9222-9227 9229-9242 9244 9257 9264 9304 9315-9316 9325-9330 9368 9401 9455 9466 9484 9525 9542 9544 9563 9646 9695 9703 9780-9781 9855 9928 9942 9947 10000 10019 10027 10103-10117 10119-10120 10173-10175 10183-10185 10346 10376 10440 10470 10496 10611 10651-10657 10666 10679 10695 10698-10704 10760 10772 10778 10782 10788 10891 11033 11051-11052 11061 11063 11112 11231 11256 11356-11357 11359-11360 11362-11367 11369-11372 11376 11385 11406 11416 11421-11425 11462 11513 11588 11605-11607 11620 11833 11869 11932 11980 11985 12006 12017 12033 12143 12160 12193 12195 12205 12207-12208 12292 12332-12336 12338-12345 12347 12349-12356 12358-12360 12363 12368 12405 12420 12422 12424 12428 12434-12439 12465 12523 12601 12605 12609-12611 12617 12636 12645 12670 12691 12707 12740 12749 12754 12768 12775 12777 12817-12824 12830 12840 12845-12848 12863 12911 12940 12955 12977 13020 13026-13028 13038 13047 13051 13062 13072-13073 13090 13104 13117 13126 13195 13202 13254 13295-13296 13360-13371 13389 13391 13399 13412 13479 13488-13489 13492 13494 13498 13500 13503-13504 13506-13507 13511-13512 13516 13533 13551 13554 13590 13613 13627 13644 13646 13656 13659-13660 13670 13713 13721 13743 13771-13774 13782 13803 13809 13812-13814 13860 13866-13867 13869 13872 13882 13884-13885 13888 13891 13893
trachea	Clontech	TRC001	83 438 483 858 1006 1034 1242-1245 1359 1483-1494 1616-1619 1621 1671 2266 2324 2388-2391 2409 2438-2439 2444 2575 3001 3136 3272 3425 3460 3535 3548 3748-3756 3810 3962-3967 4083-4086 4530 4755 4758 4949 4957 4987 5075 5213 5295-5299 5481-5482 5484 5486-5488 5490-5491 5535 5576-5580 5656 5941 6091 6450-6453 6574-6578 6637 6797

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
			6955 7041-7042 7049 7068 7204 7347 7733 7864 7894 8034-8043 8045 8120 8139 8144 8563 8635 8646 9012 9092-9099 9245-9251 9317 9331 9333 9484 9503 9517 9734 10020-10022 10027 10097 10121-10123 10125-10127 10284 10344 10478 10612 10658-10659 10705 10812 10948 11145 11233-11239 11373-11376 11428-11429 11942 12022 12209-12211 12292 12361-12363 12365-12366 12438 12440-12441 12611 12617 12734-12735 12825-12826 12900 12904-12905 12916 12923 12965 12972 12976 12989 13029-13030 13306-13308 13372-13373 13492 13670 13713 13722 13775-13781 13803 13884
uterus	Clontech	UTR001	1036 1134 1246-1250 1324 1493 1495-1511 1622 1671 1928 2145 2266 2310 2326 2343 2367 2392-2397 2555 3281 3479 3720 3757-3760 3762 3969-3986 4088-4090 4192 4201 4246 4277 4390-4391 4434 4515 4998 5002 5184 5300-5301 5481 5492-5504 5656 5695 5712 5794 5936 6116 6475 6579-6587 6595 6870 7002 7043-7044 7311 7484 7895-7899 8046-8055 8057-8059 8061 8110 8146-8147 8152 8226 8262 8622 8677 9008 9092 9137 9252-9257 9334 9370 9414 9466 9539 9703 10023-10024 10128-10134 10346 10595 10660-10668 10706-10708 11145 11239 11289 11377 11379-11385 11695 11830 12367-12369 12371-12374 12442-12443 12570 12670 12693 12827-12829 12914 13002 13047 13073 13083-13084 13131 13260 13277 13309-13310 13374-13376 13394 13489 13512 13713 13782 13784-13787 13866 13869

TABLE 2

SEQ ID NO:	Accession No.	Species	Description	Smith-Waterman Score	% Identity
1	M36501	Homo sapiens	alpha-2-macroglobulin	118	69
2	AF118090	Homo sapiens	PRO2044	247	59
3	X01683	Homo sapiens	alpha 1-antitrypsin	544	78
4	L27428	Homo sapiens	reverse transcriptase	79	27
5	M22332	Homo sapiens	unknown protein	89	40
6	AF015539	Mytilus edulis	precollagen P	113	33
7	X03325	Homo sapiens	apolipoprotein B fragment	540	83
8	AB019280	Mus musculus	sprouty-4	91	35
9	D88152	Homo sapiens	acetyl-coenzyme A transporter	625	87
10	G00328	Homo sapiens	Human secreted protein, SEQ ID NO: 4409.	124	58
11	AL049569	Homo sapiens	dJ37C10.5 (KIAA0445)	182	82
12	AJ242540	Volvox carteri f. nagariensis	hydroxyproline-rich glycoprotein DZ-HRGP	85	37
13	L27428	Homo sapiens	reverse transcriptase	135	61
14	U49973	Homo sapiens	ORF1; MER37; putative transposase similar to pogo element	266	72
15	U93569	Homo sapiens	putative p150	135	37
16	G02832	Homo sapiens	Human secreted protein, SEQ ID NO: 6913.	132	67
17	X53581	Rattus norvegicus	ORF4	124	34
18	AF183961	Homo sapiens	carbon catabolite repression 4 protein homolog	431	75
19	AJ002190	Homo sapiens	dihydroxyacetone phosphate acyltransferase	551	88
20	Y12713	Mus musculus	Pro-Pol-dUTPase polyprotein	127	45
21	AK001269	Homo sapiens	unnamed protein product	1643	99
22	U49973	Homo sapiens	ORF1; MER37; putative transposase similar to pogo element	275	58
23	G03793	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	123	75
24	AF156550	Mus musculus	putative E1-E2 ATPase	168	58
25	AF119856	Homo sapiens	PRO1851	585	83
26	U49974	Homo sapiens	mariner transposase	187	46
27	G00901	Homo sapiens	Human secreted protein, SEQ ID NO: 4982.	86	30
28	AF295773	Homo sapiens	ral guanine nucleotide dissociation stimulator	126	74
29	AF113685	Homo sapiens	PRO0974	92	73
30	U83303	Homo sapiens	line-1 reverse transcriptase	102	50
31	Y91577	Homo sapiens	Human secreted protein sequence encoded by gene 2 SEQ ID NO:250.	279	75
32	AF003535	Homo sapiens	ORF2-like protein	114	47
33	M15386	Homo sapiens	gamma-globin	370	84
34	M19419	Mus musculus	proline-rich salivary protein	110	35
35	AF211943	Homo sapiens	WW domain-containing protein WWOX	586	83
36	X13885	Nicotiana tabacum	extensin (AA 1-620)	103	35
37	U93563	Homo sapiens	putative p150	127	58
38	U93564	Homo sapiens	putative p150	103	77
39	AF069732	Homo sapiens	ADA2-like protein	524	88
40	X61046	Hydra sp.	mini-collagen	101	34
41	AK000322	Homo sapiens	unnamed protein product	566	80
42	G03646	Homo sapiens	Human secreted protein, SEQ ID NO:	103	57

SEQ ID NO:	Accession No.	Species	Description	Smith-Waterman Score	% Identity
5096	R13556	Homo sapiens	Protein encoded downstream of hhc_M oncoprotein.	129	67
5097	U64849	Caenorhabditis elegans	Contains similarity to Pfam domain: PF00646 (F-box), Score=28.7, E-value=4.3e-05, N=1	197	39
5098	AF006010	Homo sapiens	progestin induced protein	174	86
5099	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	119	55
5100	G03043	Homo sapiens	Human secreted protein, SEQ ID NO: 7124.	132	82
5101	G00403	Homo sapiens	Human secreted protein, SEQ ID NO: 4484.	159	77
5102	S58722	Homo sapiens	X-linked retinopathy protein {C-terminal, clone XEH.8c}	99	70
5103	J02459	bacteriophage lambda	J (tail:host specificity;1132)	485	96
5104	J02459	bacteriophage lambda	H (tail component;853)	823	98
5105	AF025467	Caenorhabditis elegans	contains similarity to drosophila DNA-binding protein K10 (NID:g8148)	106	42
5106	AF130089	Homo sapiens	PRO2550	121	35
5107	G04078	Homo sapiens	Human secreted protein, SEQ ID NO: 8159.	135	81
5108	P92219	Homo sapiens (human)	CR1 protein.	123	88
5109	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	83	78
5110	U41751	Mus musculus	EI24	560	83
5111	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	150	69
5112	AF130089	Homo sapiens	PRO2550	164	88
5113	G03084	Homo sapiens	Human secreted protein, SEQ ID NO: 7165.	160	72
5114	AF052831	Trypanosoma cruzi	unknown	107	70
5115	AL359782	Trypanosoma brucei	possible (hhv-6) u1102, variant a dna; complete virion genome.	128	68
5116	X55684	Lycopersicon esculentum	extensin (class I)	60	29
5117	W80406	Homo sapiens	A secreted protein encoded by clone dh40_3.	124	64
5118	AP000616	Oryza sativa	similar to RING-H2 finger protein RHA1a (AF078683)	115	60
5119	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	79
5120	X65718	Prunus dulcis	extensin	94	33
5121	AP000616	Oryza sativa	similar to RING-H2 finger protein RHA1a (AF078683)	132	100
5122	AJ007714	Homo sapiens	lysine-ketoglutarate reductase /saccharopine dehydrogenase	154	96
5123	AJ007714	Homo sapiens	lysine-ketoglutarate reductase /saccharopine dehydrogenase	158	80
5124	AL359782	Trypanosoma brucei	probable similar to ring-h2 finger protein rha1a.	109	44
5125	AF130089	Homo sapiens	PRO2550	112	46
5126	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	129	60
5127	G00442	Homo sapiens	Human secreted protein, SEQ ID NO: 4523.	99	68

TABLE 3

SEQ ID NO: of nucleotide sequence	SEQ ID NO: of peptide sequence	Method	SEQ ID NO: in USSN 09/515,126	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1	13902	A	1	114	434	AIFKCVEGMFRIAMVNVCFVSSGSLLI* PLTY/GVYDEWTHFAYMTIDLLEIPITG SHPVVLNALFCLEAP\WISPNTGSFAYP VYPKSLIAHDFAVEATMPYIRLSST
2	13903	A	2	124	466	KSNIPNLGDCGWESLFNR\QSWRSSLAV \NDTYSSKKSNAETFTFHADLCTLSDKD RPITIQTALAELVKHKPKATYEQLIAVL DEF/ANFLKKWWKAYDKENLFCEEG*KL CAASN
3	13904	A	4	1	427	EGFLELLRTRNHSNSQLQLTTGIGLFLN EGLKLVDKFLEDV*K*YHSETFTVNFSD TE*AMKHINDYVEKGTQGKIVDLVKELD RDTVFDLANYIFFKGKWDRPFEVNDTEE EDFHVDQVSTVNEPIMKLLS\MLNIHPC FKL
4	13905	A	5	1	464	KIKSFYASKDTIKRMRVTDWKKIFAY*I SDKELIFTLRTLKT**K*GKQPNLKNGQ EI*VPISPQDIQIAHK\HLEGWSTSLVS ELAPCEAPV/RHPLTGLTIAGLQGFGEA GRLVRGRWGC*W/VHPF*KYI/WQFLSK LHISLPYDPTTPLLGTCSR
5	13906	A	6	308	3	HFVIHSKHDLAIAHLGIY/PREMKT*VH TKTCT*IFTVALSVIARAWNQPGRPLCS EWL\KYMVHTME*HSAIKRLNYRYKNNC VNLFLGITLSEKSQTQNVI
6	13907	A	7	587	2	FLTRETGDPTGRSSSHANTQSRFFPDDP PG\PLNNLGNTHGCGRRAGRCPGTGPDG P\AGCGGPRCWPSGHLAATGD*GPSCGR LGANRGEAGPAGFTACSPLSGCRTPYTH HFPASRMSCHLNCASPRTYRSQGNRGCE RVAQGSQGAGGERGAKSQVPVPAPARNK DPAKCRKPRNRRPGNSGPVVRAYRRQR
7	13908	A	8	1	474	RILNEEHGKYEGLHE*EVKWHLYIKSPA FTDLHLCYQKDMNGISTSASSPAVGTVG MDMDEDDDFSKWNFYYSPHSYPDK*LTI FKTESRVRESDEVTQIKVNWDEEVISGL LTSLKDNVLKATGVLYDYAYK\YLCEHT RSTLKEESLKLERNLQNH
8	13909	A	9	3	539	SQCSPFISPACSLTALEEETEALRVHPR LCLSPNLAPSSGPPRPPELAPCPPSSQA GLRTCHSWVKGLHQPLPVASGMKSTFCN KTYTCPYPPP/PPLCPNHSPNALTLPDS VTHAVPPE/L*SPSAPPSSTA*ILGSPS \CGASPCNHPHSHPGICPTPPGLWPVCP CAPRAWQRDGTRQT
9	13910	A	10	2	453	RL*LGLEYALLVWGTPKV*H*GGFPIYY YIVLLLSYALHQVTEYSMYVSIMAFNAK VSDPLIVGTYMTLLNTVSNLIGNWPSTV SLWLVNPLTVKECV*TSYQNCCTPDAAE LCKKLGGSCVTALDGYYVESIICVSIAF V/W*VFLVHKFK
10	13911	A	11	20	475	KMGVPPLLMSDPNRFLFPKNFLREKTIS PPKTF*PLKIWVKGQWVLNFLGFPGFKI FFPVFKFFFFFFFF/RDRVSLYHPGWSA VSQSELTAALTSPGSGDQVILPSQPPK* ENHLNLGGRGCSEPRLPRAEFLDLRSFS

SEQ ID NO: of nucleotide sequence	SEQ ID NO: of peptide sequence	Method	SEQ ID NO: in USSN 09/515,126	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						RDHPGQHGETPYPL
5110	19011	A	5147	3	413	SFYRGFIPVLQTVTVLIIGDPSLHGDAW SWMEFFLTVIFIALWMLPLIGLTIVVNA IWFQDIADLAFEVSWTKPHPFPSVSQII ADMLFPLLLQALFLIQGIAVSLFPIHLV GQLDSLLHMTLLYSLYCSEYRLVIN
5111	19012	A	5148	409	247	QTKFRHVAQAGPHFLGSSDVPTSASQSV GITGMSLHAQHQHFFTASMGRAFGEN
5112	19013	A	5149	284	1	AEASRAHMEAHARTQGTPGSKAEGVGPV EEYRIWRRPGPDRAHASDSEESGEPGGP PPHPANFVFLVETGFFRVGQAGLEFPTS GDPAASASPV
5113	19014	A	5150	439	120	FGSPRLECSGVISAHCNLRLLGSSDSPA SASHVAGITARPNARIIKAFGAPSTCTV VWERHEQQTCIQRSDLPAGRQKKTGPKN RPGLVAFPFNASPLGGPGGRIA
5114	19015	A	5151	415	0	FFFFFFFFFFFFLPFVQSFYIYIYMYI
5115	19016	A	5152	67	179	ICLSLFLTVHVCVCVCVCVCVCVCVCEM VFGVSIPCN
5116	19017	A	5153	280	2	KTPFFLNLKEKTRAPFFLKKTFFKKKPF KIFFSKKRNFFFFPPPGKRPPTLFKKKF FFFFQKKKNFFFSPKKKKKKKKKGRRS RSRTSPRV
5117	19018	A	5154	296	406	DRVLLCHSSCSEVAESRLTAASTSWAQA ILLPQPPEW
5118	19019	A	5155	114	402	VWRIKASVEKKKKKKKKKQKKKKKKKR
5119	19020	A	5156	110	1	QRFLLTLKSLILRLGAVAHACNPSTLGG QGGRIMRS
5120	19021	A	5157	382	95	SSSSPHFLTPSQLRLFFPFFPLKIFFFP KAFNFCGGVFPIFSPPKKKFFFQNSQVG FKNPPQKEKNFSFPPPVKFGPPKGFFKR PPPFFFFFFFFF
5121	19022	A	5158	2	350	TLQPGRQSETPFQKKKKKKKKKKKKKKK KKKKKKRGGGFLKNLGGGPKIWGVEKFF FFFFGGGYKKPLGFFLEKIFFLGGGIWA PPPPKKISFWGKKKFFWGGGGKTPLFFC GGKK
5122	19023	A	5159	1	208	PTRTLYLVKLGLMGPAEIFLSSFQEIRL IDYDKMVDHRGVRVVAFGQWAGVAGKYA RAYITHFLNWLGE
5123	19024	A	5160	1	208	PTRPLYLLKLSLMGPVEIFLSSFQEIRL IDYEKMVDHRGVRVVAFGQWAGVAGKYA RAYITHFLNWLGE
5124	19025	A	5161	419	25	AQKKKKKKKKKKKKKGGGPLKKNFGGGK NKRGKKKKNFFLKRGGKKKTGGILEKKN FFGGGKKGEKPPQKKKALKEKKKFLRGK GGKKPLNWGGKKKW
5125	19026	A	5162	475	150	KPRTSGTARVPGDPPPRCLDDKRLPSRP VIFVFLAETECDRVGQDGLQLLTSGDHP HLTFPNCYNYRREPPRPTSSGFQHDIET LHYRRGKADRFKAHLPK
5126	19027	A	5163	237	409	FHNQKTYLDSALCTLVYFVVFVLRDGIF LLPQLECSGTVMAHCNLKLLGSRSPPAL A
5127	19028	A	5164	373	476	LGFILLSFLETRSHSVAQAEVQWCDHGL LQPQTP
5128	19029	A	5165	374	132	IFYSPFLYYHMLGKYHKIKNLKIKHMWL DAVAHACNPSTLGGRGGRITRLGDRDHP

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-13901 , a mature protein coding portion of SEQ ID NO: 1-13901 , an active domain of SEQ ID NO: 1-13901 , and complementary sequences thereof.

2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.

3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.

4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.

5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.

6. A vector comprising the polynucleotide of claim 1.

7. An expression vector comprising the polynucleotide of claim 1.

8. A host cell genetically engineered to comprise the polynucleotide of claim 1.

9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.

10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:

(a) a polypeptide encoded by any one of the polynucleotides of claim 1; and

(b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NOS: 1-13901.

11. A composition comprising the polypeptide of claim 10 and a carrier.

12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and

b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.

14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;

b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and

c) detecting said product and thereby the polynucleotide of claim 1 in the sample.

15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.

16. A method for detecting the polypeptide of claim 10 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and

b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.

17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and

b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

19. A method of producing the polypeptide of claim 10, comprising,

a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-13901 , a mature protein coding portion of SEQ ID NO: 1-13901 , an active domain of SEQ ID NO: 1-13901 , complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-13901 , under conditions sufficient to express the polypeptide in said cell; and

b) isolating the polypeptide from the cell culture or cells of step (a).

20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 13902-27802, the mature protein portion thereof, or the active domain thereof.

21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.

22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NOS: 1-13901.

23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.

24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.

25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.

26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.